Humboldt-Universität zu Berlin – Geographisches Institut

Quantifying urban land cover by means of machine learning and imaging spectrometer data at multiple spatial scales

DISSERTATION

Zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)

Im Fach Geographie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Dipl. Geogr. Akpona Okujeni

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter:
Prof. Dr. Patrick Hostert
Prof. Dr. Ben Somers
Prof. Dr. Björn Waske

Eingereicht: 20. August 2014
Tag der Verteidigung: 21. November 2014

Acknowledgments

Completing this doctoral thesis would not have been possible without the support and encouragement of many kind people around me to whom I am very grateful.

First and foremost, I would like to express my gratitude to the two persons who introduced me into the world of remote sensing and ever since supported my scientific career in every possible way: *Sebastian van der Linden*, thank you for not only being an excellent and inspiring supervisor, but also for the trust and outstanding support I have received during the last years. This is exceptional and I deeply appreciate that! *Patrick Hostert*, thank you for not only being an excellent and inspiring supervisor too, but also for being such a great head of the Geomatics Lab. The research environment you have established is unique and I feel truly privileged to be a part of your team!

Second, I am grateful to my colleagues and friends at the *Geomatics Lab*, the *Biogeography and Conservation Biology Lab*, and the *Geoinformation Science Lab*. We are sooo many now, that's why I thank *ALL* for your support, contributions and great company throughout the last years! Nevertheless, I would like to point out my two long-time companions *Patrick Griffiths* & *StefanSüß* – indeed, "we have really come a long way now…". *Benjamin Jakimow*, *Andreas Rabe*, *Marcel Schwieder* and *Maria Cierpinski*, many thanks for directly contributing to this thesis in one way or the other. *Tobia Lakes*, thank you for being head of the committee. *Anika Sieber*, thank you for being there whenever needed! *Dagmar Wörister*, thank you for all your helping hands! Without your reliable and constant support, focusing mainly on research would not have been possible.

Third, several scientists crossed my way during the period being a doctoral researcher and I particularly thank *Ben Somers* and *Björn Waske* for being external supervisors of this thesis and for the nice conversations beyond remote sensing we had when met in Berlin, Leuven, or elsewhere. Thank you *Laurent Tits* and *Jochem Verrelst* for contributing as co-authors to this thesis. A big thank also to *Christopher Small* and the *German Academic Exchange Service* for a memorable research stay at the Columbia University in New York.

On the personal side, I am thankful to all my *friends* and *soulmates*, who are cycling around the world or who I have heavily neglected during the past months.

Last but not least, I would like to dedicate this thesis to my *family* and express my heartfelt gratitude for their continuous support, trust and love.

Abstract

The global dimension of urbanization constitutes a great environmental challenge for the 21st century. Remote sensing is a valuable Earth observation tool, which helps to better understand this dynamic process and its ecological implications. The scientific focus of this work was to quantify urban land cover by means of machine learning and imaging spectrometer data at multiple spatial scales. Experiments considered innovative methodological developments and novel opportunities in urban research that will be created by upcoming hyperspectral satellite missions, specifically the Environmental Mapping and Analysis Program (EnMAP). Airborne Hyperspectral Mapper data at 3.6 m and 9 m resolution and simulated spaceborne EnMAP data at 30 m resolution were used to map land cover along an urban-rural gradient of Berlin, Germany. In the first part of this work, the combination of support vector regression with synthetically mixed training data was introduced as sub-pixel mapping technique. Results demonstrate that the approach performs well in quantifying thematically meaningful yet spectrally challenging surface types such as rooftops and paved-areas, or grass and tree cover. The method proves to be both superior to other advanced regression algorithms and traditional multiple endmember spectral mixture analysis and universally applicable when proceeding from the spatial scale of airborne to spaceborne hyperspectral data. In the second part of this work, the value of future EnMAP data for urban remote sensing was evaluated. Detailed explorations on simulated data demonstrate their suitability for improving and extending the approved Vegetation-Impervious-Soil mapping scheme. Comprehensive analyses of benefits and limitations of EnMAP data reveal both challenges caused by the high numbers of mixed pixels, when compared to hyperspectral airborne imagery, and improvements due to the greater material discrimination capability, when compared to multispectral spaceborne imagery. In summary, findings demonstrate how combining spaceborne imaging spectrometry and machine learning techniques could introduce a new quality to the field of urban remote sensing. This is of importance, especially considering the need for reliable global urban land cover products.

Zusammenfassung

Das weltweite Ausmaß der Urbanisierung zählt zu den großen ökologischen Herausforderungen des 21. Jahrhunderts. Als eine Form der Erdbeobachtung bietet die Fernerkundung die ideale Möglichkeit das Verständnis dieses dynamischen Prozesses und seiner Auswirkungen zu erweitern. Der wissenschaftliche Schwerpunkt dieser Arbeit lag in der Quantifizierung der städtischen Landbedeckung mittels Maschinellen Lernens und räumlich unterschiedlich aufgelöster Hyperspektraldaten. Die Untersuchungen berücksichtigten innovative methodische Entwicklungen und neue Möglichkeiten für die Stadtforschung, die durch die bevorstehende hyperspektrale Satellitenmission *Environmental Mapping and Analysis Program* (EnMAP) geschaffen werden. Auf Basis von Bilddaten des flugzeugestützten *Hyperspectral Mapper* Sensors mit einer Auflösung von 3,6 m und 9 m sowie simulierten EnMAP-Satellitendaten mit einer Auflösung von 30 m wurde eine Landbedeckungskartierung entlang des Stadt-Umland-Gradienten Berlins, Deutschland, durchgeführt. Im ersten Teil der Arbeit wurde die Kombination von *Support Vector Regression* mit synthetisch gemischten Trainingsdaten für die Subpixelkartierung eingeführt. Die Ergebnisse zeigen, dass sich der Ansatz gut zur Quantifizierung thematisch relevanter und spektral anspruchsvoller Oberflächenarten, z.B. bebaut und unbebaut versiegelte Flächen oder Grasflächen und Baumbestände, eignet. Die Methode liefert verbesserte Ergebnisse gegenüber weiteren fortgeschrittenen Regressionsalgorithmen und dem gängigen Ansatz der multiplen Entmischungsanalyse und zeigt sich als universell einsetzbar bei der Überbrückung des räumlichen Skalensprungs von flugzeuggestützten hin zu satellitengestützten Hyperspektraldaten. Im zweiten Teil der Arbeit wurde der Wert der zukünftigen EnMAP-Daten für die städtische Fernerkundung abgeschätzt. Detaillierte Untersuchungen mittels simulierten Bilddaten unterstreichen ihre Eignung für eine verbesserte und erweiterte Beschreibung der Stadt nach dem bewährten *Vegetation-Impervious-Soil*-Schema. Umfassende Analysen der Möglichkeiten und Grenzen von EnMAP-Daten zeigen sowohl Nachteile durch die höhere Anzahl von Mischpixel im Vergleich zu hyperspektralen Flugzeugdaten als auch Vorteile aufgrund der verbesserten Differenzierung städtischer Materialien im Vergleich zu multispektralen Satellitendaten. Insgesamt veranschaulichen die Ergebnisse dieser Arbeit, dass die Kombination von hyperspektraler Satellitenbildfernerkundung mit Methoden des Maschinellen Lernens eine neue Qualität in die städtische Fernerkundung bringen kann. Dies ist von großer Bedeutung, insbesondere vor dem Hintergrund des Bedarfs an verlässlichen Datenprodukten der globalen städtischen Landbedeckung.

Contents

List of Figures

List of Tables

Chapter I:
Introduction

1 Scientific background

1.1 The urban world

The cultivation of land is one of the fundamental characteristics of civilization on Earth (Vitousek et al. 1997; DeFries et al. 2004; MEA 2005). Areas transformed for urban living occupy only a small percentage of the Earth's land surface (DeFries et al. 2004; Kareiva et al. 2007). For this reason, changes urban land area per se do not appear to be central in land cover change worldwide (Lambin et al. 2001). Yet, urbanization is a dominant demographic, socio-economic, and biophysical process, and a major contributor to global environmental change (Lambin et al. 2001; Grimm et al. 2008; Pickett et al. 2011).

The 21st century is often referred to as the "Century of the City" (Seto et al. 2010). In 1900, only 13% of people lived in cities. Today, there are 3.6 billion urban dwellers, making up more than 50% of the world's population. Since 1950, the number of cities with one million or more citizens has increased from 75 to 450. Since 1970, another 21 megacities with more than ten million inhabitants have emerged next to New York and Tokyo (United Nations 2006, 2012a, 2012b). Given the projected gains in global population by 2.3 billion and urban population by 2.6 billion until 2050 (United Nations 2012a), it is obvious that the world is becoming increasingly urbanized. High growth rates are particularly expected in the developing world and fast-developing countries of Asia, where the percentage of urban population is mostly far below the 75-80% urbanization level present in the more developed regions (Anderson et al. 2013; Kronenberg et al. 2013; McPhearson et al. 2013; Pauchard et al. 2013; Seto 2013). China and India in particular are predicted future hotspots of global population growth and urban development (United Nations 2012a; Fragkias et al. 2013).

The sheer quantity of urban growth is one reason to study urban areas. Another compelling argument is the quality of the urbanization process and its broader implications. On the one hand, urbanization has many positive effects. For example, cities are considered as centers of economic growth and better living conditions (Quigley 1998), as gateways to the world's economy (Sánchez-Rodríguez et al. 2005), as hubs of productivity, employment, innovation, and creativity (UN Habitat 2010), as places where political, social, and cultural life takes place (Satterthwaite 2007), and so forth. This point of view is often underlined by the fact that no country has ever achieved sustained economic or social development

without being urbanized (UN Habitat 2010). On the other hand, urbanization is one of the most powerful, irreversible, and visible anthropogenic forces on Earth, leading to severe environmental and social problems (Sánchez-Rodríguez et al. 2005). Urban living depends on the resources of ecosystems far beyond city boundaries, which leads to extraordinary large ecological footprints (Rees 1992; Folke et al. 1997). The spread of impervious areas, the loss of natural environments, and increased landscape fragmentation along with the concentration of human consumption, production, and waste generation have altered biogeochemical cycles, climate regimes, hydrological systems, or biodiversity at multiple scales (Grimm et al. 2008). Environmental impacts are global considering that urban areas are, for example, major sources of CO_2 and other greenhouse gases, which considerably contribute to global climate change (Kennedy et al. 2009). At the regional and local scale, phenomena such as the urban heat island (Oke 1982; Grimmond 2007), air pollution (Fenger 1999), or the increased vulnerability to natural hazards (Pelling 2003) are prominent characteristics of urban environments. Such changes in ecosystem condition and functioning can, in turn, influence the quality of life and human health (Harlan et al. 2006; Tan et al. 2010). Not least, cities especially in less the developed world are subject to fast-sprawling informal settlements and severe socio-economic disparities leading to what is referred to as the "urban divide" (UN Habitat 2010).

Meeting fundamental human needs while preserving life support systems of the Earth is a key effort to create a more sustainable world (NRC 1999). Given the global dimension and implications of urbanization, any efforts seeking to promote sustainable and desirable future developments have to consider urban dynamics for understanding and managing the fast changing world (Herold 2009). Policies must be adapted to sustain benefits of urbanization and to anticipate and manage negative consequences of urban growth (UN Habitat 2010). Contemporary urbanization has the chance to promote ecological sustainability due to increased environmental awareness and returns from innovation, productivity, and efficiency (Grimm et al. 2008; Seto et al. 2010). Given this perspective, interactions between socio-economic and environmental processes in urban landscapes need to be better understood (Alberti 2005). Integrated research of the diverse aspects and implications of urban systems in times of rapid urbanization is of great importance to face the complexity of the involved dynamics. This may ultimately support decision makers involved in governance, planning, or conservation to promote the sustainability of existing or emerging cities (Grimm et al. 2008; Seto et al. 2010; Pickett et al. 2011).

1.2 Earth observation of urban areas

Earth observation (EO) data is essential to understanding of natural and human-induced changes and their implications within the Earth's spheres. Furthermore, EO data is essential to build information products, forecast models, and other tools for making informed decisions. In this context, satellite data have played a key role in studying, for example, stratospheric ozone levels, air pollution, glacial and sea ice contingents, land-use changes, weather patterns, ocean circulations, or natural disasters (NRC 2007).

By providing synoptic views in space and time, remote sensing has become a valuable EO tool to measure and monitor the Earth's environment (Chuvieco 2008; Liang 2008; Weng 2011). As one area of application, urban remote sensing has emerged as new frontier in EO technology (Weng 2014) and holds the potential to link research, policy, and practice for successfully solving pressing urban and global change issues (Seto 2009). For example, Jensen and Cowen (1999) and Heiden et al. (2012) propose the use of remote sensing data as alternative to cost- and time-intensive field surveys for building inventories of urban surface properties required for urban planning. Masser (2001) and Maktav et al. (2005) stress the value of remote sensing to derive up-to-date spatial information on urban areas in the developing world, where alternative sources of information are limited. Miller and Small (2003) mention the opportunity to identify, monitor, and apprehend a number of urban environmental problems by means of high resolution imagery from EO satellites, recent advances in global connectivity, information technology, and other existing data sources. Accordingly, all authors underscore the value of remote sensing for studying urban systems in times of rapid urbanization.

With the beginning of modern EO in the 1970s, a multitude of spaceborne sensors, including optical, thermal, radar, or night light imaging systems, have provided different types of image data globally. Ever since, most urban image analyses have been focusing on characterizing urban environments at local to regional scales (Herold 2009). Studies encompass the monitoring of urban growth (Griffiths et al. 2010; Bagan and Yamagata 2012), the assessment of urban land cover (Powell et al. 2007; Myint et al. 2011), or the quantification of environmental conditions and biophysical processes such as the urban heat island or urban forest carbon storage (Myeong et al. 2006; Weng 2009). Global urban mapping efforts have so far focused on delineating the extent of human settlements by primarily distinguishing "urban land" from "non-urban environments". Potere et al. (2009) and Schneider et al. (2010) list and discuss ten global maps of urban extent, most of them

based on the direct or indirect input of low (>100 m) spatial resolution data acquired by spaceborne multispectral sensors (e.g., MODIS, SPOT-Vegetation, or MERIS) or by night light systems (e.g., DMSP-OLS). However, large differences between urban extent estimates are reported (Potere et al. 2009), and maps are provided at very coarse spatial resolution for single time steps (Taubenböck et al. 2012). A recent improvement is the Global Urban Footprint product of the German Aerospace Center (Figure I-1). This product has been developed at 12 m resolution by means of TerraSAR-X and TanDEM-X radar data (Esch et al. 2012; Esch et al. 2013). The opening of the Landsat archive has facilitated other recent developments in global urban mapping: The continuous record of Landsat data at 30 m resolution allows detailed global mapping of urban extent over time. These data have been, for example, exploited by Taubenböck et al. (2012), who has mapped almost 40 years of megacity growth worldwide.

The proliferation of various sources of satellite imagery and processing and modeling tools has created the opportunity for progressing toward global observations of urban patterns and dynamics (Herold 2009). While the demographic shift toward an increasingly urbanized world is well documented, the physical component of urbanization at global scale is not well-understood (Seto et al. 2010). However, characterizing urban areas not only by extent, but also in terms of their physical properties could significantly advance our understanding of environmental processes related to urban meteorology, hydrology, and ecology (Small 2009). The derivation of reliable and standardized global urban land cover products is therefore one major task for urban remote sensing.

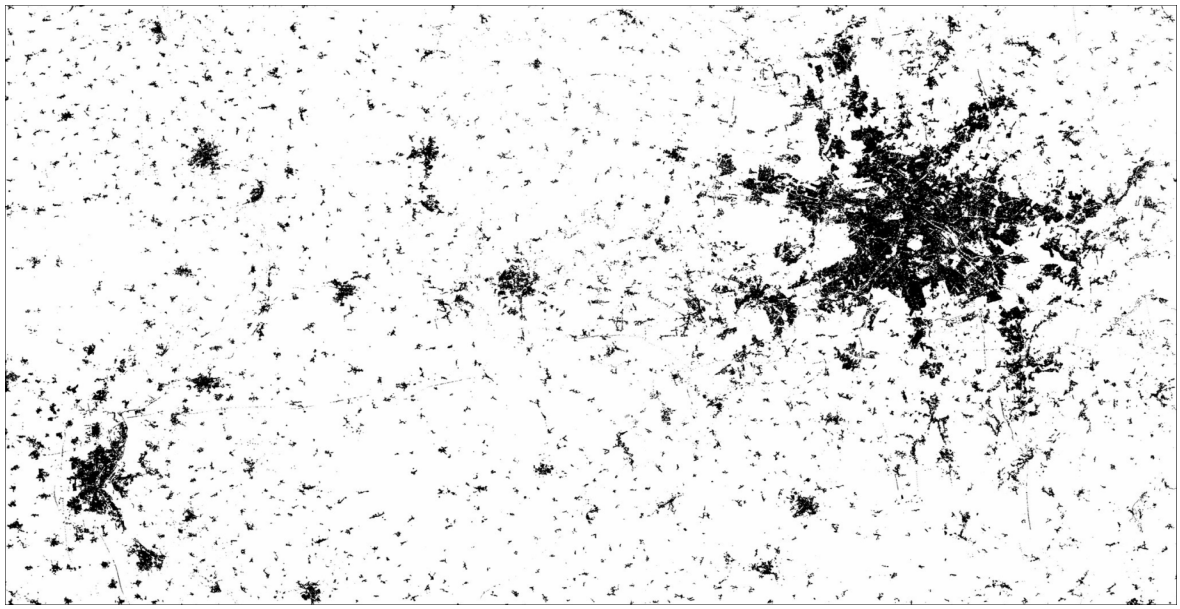


Figure I-1: Berlin-Brandenburg region as captured by the Global Urban Footprint product of the German Aerospace Center (Source: www.dlr.de).

1.3 Remote sensing-based urban land cover mapping

The use of remote sensing data for deriving descriptions of urban land cover was demonstrated in a multitude of case studies. Most mapping assessments exploited imagery from spaceborne multispectral sensors, which provide few broadband reflectance measurements in the visible to shortwave-infrared region of the electromagnetic spectrum. Medium (10-100 m) spatial resolution data (e.g., Landsat, or ASTER) were used to quantify urban land cover into the broad categories “impervious”, ”vegetation”, and “soil” (Powell et al. 2007; Pu et al. 2008) according to Ridd’s (1995) VIS framework. Very high to high (<10 m) spatial resolution imagery (e.g., WorldView 2, QuickBird, or IKONOS) were adopted to map more detailed surface properties, including different impervious and vegetation types (Myint et al. 2011; Pu and Landry 2012). Although these data sets are globally available, little progress has yet been made toward a global map of urban land cover. On the one hand, challenges due to the spatial complexity (Pesaresi and Ehrlich 2009) and limitations due to cloud coverage, data storage issues, data costs, and high processing efforts (Taubenböck et al. 2012) are reported for higher spatial resolution data; medium spatial resolution urban data face the challenge of large numbers of mixed pixels (Small 2001; Powell et al. 2007). On the other hand, general drawbacks of multispectral data are frequently mentioned when mapping urban land cover on a purely spectral basis. Many urban materials are not spectrally distinct when represented by broadband reflectance spectra, which can lead to considerable confusion between various impervious surfaces and bare soil, or between different vegetation types (Thomas et al. 2003; Powell et al. 2007; Myint et al. 2011). The SVD model (Small and Lu 2006) translates the spectral mixing space into the biophysical surface reflectance properties “substrate”, “vegetation”, and “dark matter”. This approach avoids spectral ambiguities and at the same time offers a robust way to characterize urban areas worldwide in terms of their physical properties (Small 2003, 2009). Although SVD was used to derive meaningful estimates of urban vegetation (Small and Lu 2006), linking spectral with more thematically-oriented information remains desirable. Other studies account for limitations of multispectral data for urban mapping by multi-date and multi-sensor approaches (Yuan et al. 2005; Griffiths et al. 2010; Lu et al. 2011; Tigges et al. 2013).

The increasing use of hyperspectral data – also referred to as imaging spectrometer data – has advanced urban land cover mapping on a purely spectral basis. Full range hyperspectral sensors provide hundreds of contiguous spectral measurements in the visible to shortwave-infrared domain. Compared to multispectral systems, hyperspectral sensors

uncover material characteristic reflectance signatures, including specific absorption features (Goetz 2009) (Figure I-2). This property has introduced new quality into urban remote sensing. For example, Herold et al. (2004) and Heiden et al. (2007) analyzed comprehensive urban spectral libraries of anthropogenic and natural materials and point to the value of the high spectral information content for improved separability. Herold et al. (2003) and Gamba and Dell'Acqua (2006) highlight the value of hyperspectral compared to multispectral data for improved urban mapping. Studies by Roessner et al. (2001), Herold et al. (2003), van der Linden and Hostert (2009) and Franke et al. (2009) demonstrate the value of imaging spectrometer data for mapping thematically meaningful urban surface types (e.g., rooftops and paved-areas, bare soil, grass- and tree-cover), or for mapping anthropogenic materials and vegetation species.

However, the use of imaging spectrometer data for urban mapping has thus far been more experimental by nature. High quality images are only available from episodic flight campaigns, where airborne hyperspectral sensors like the Advanced Visible Near Infrared Imaging Spectrometer (AVIRIS; Green et al. 1998) or the Hyperspectral Mapper

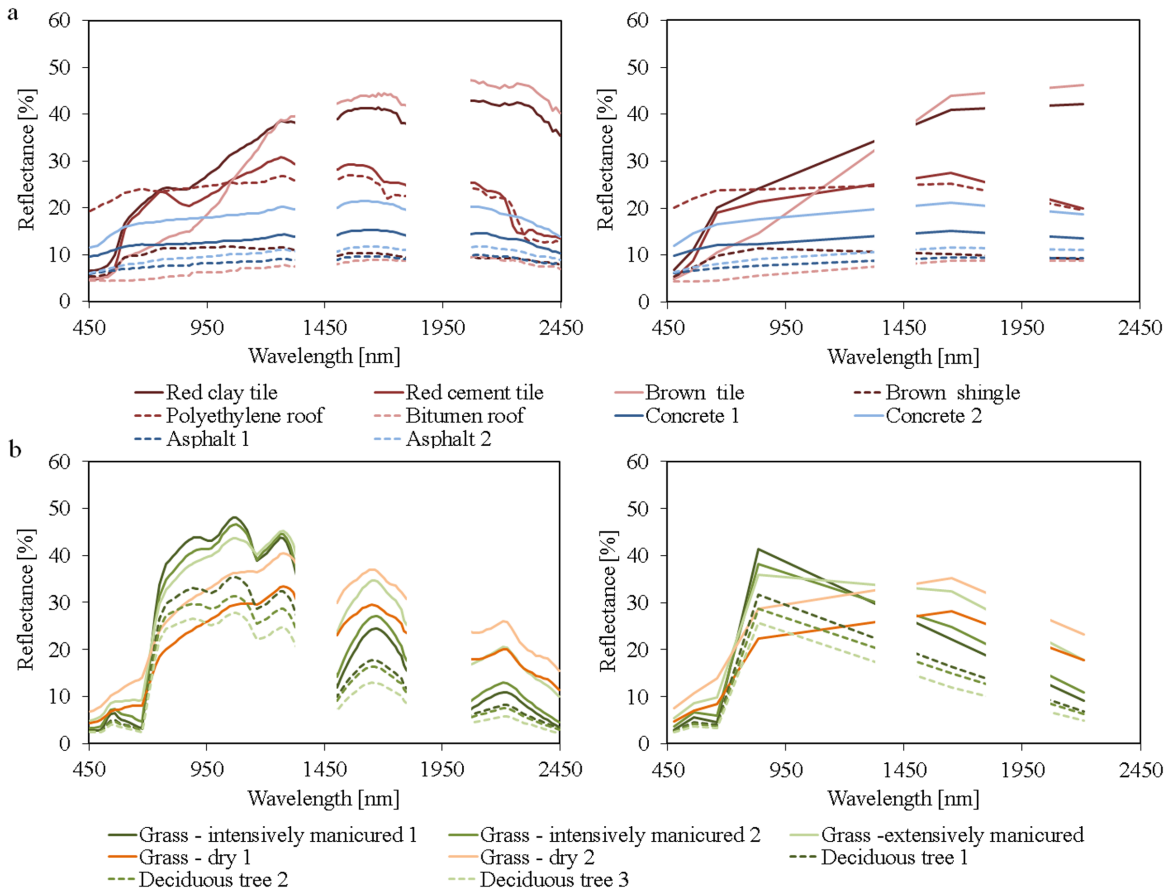


Figure I-2: Spectral library of (a) anthropogenic and (b) natural urban surface materials as recorded by HyMap and Landsat. HyMap spectra (111 bands; out of original 128 bands) provide considerable more information about overall shape and absorption features compared to Landsat spectra (6 bands; spectrally resampled from HyMap using the Landsat ETM+ filter function).

(HyMap; Cocks et al. 1998) generally capture a spatially limited subset of an urban area at high spatial resolution. Globally available data from Hyperion, which is currently the only spaceborne hyperspectral sensor, are used in a few urban studies (Cavalli et al. 2008; Weng et al. 2008). To some extent, this relates to limited instrument performances compared to the dominating airborne hyperspectral systems (Kruse et al. 2003; Roberts et al. 2003).

This lack in data availability will change soon: several operational spaceborne hyperspectral missions, which will provide high quality imaging spectrometer data globally, are currently underway (Staenz and Held 2012). This will create unprecedented opportunities for novel applications, including the global mapping of urban land cover by means of hyperspectral imagery. The Environmental Mapping and Analysis Program (EnMAP; Kaufmann et al. 2008) in particular holds great potential for urban environments (Heldens et al. 2011). With a spatial resolution of 30 m, 244 spectral bands between 420 and 2,450 nm, and a nominal revisit time of 21 days, EnMAP represents a promising hyperspectral complement to multispectral spaceborne systems with medium spatial resolution such as Landsat. Exploring the potential of upcoming hyperspectral satellite missions, such as EnMAP for urban mapping, is therefore of great importance, particularly with regard to the need for reliable global urban land cover products.

1.4 Machine learning algorithms for urban land cover mapping

The development of adequate analysis techniques to exploit the ever-expanding bulk of image data is an important research field in remote sensing. The diversity of applications and the increasing dimensionality of the data have called for powerful methods that make best use of the latent information content. Against this background, machine learning algorithms have been successfully introduced into remote sensing data analysis.

In the context of land cover mapping by means of imaging spectrometer data, urban areas are probably some of the most challenging environments to be analyzed. They are spectrally and spatially complex and require special attention when being mapped. On the one hand, urban land cover assessments face challenges such as high within-class variability and high inter-class similarity. High within-class variability is characteristic for surfaces with high spectral diversity and variability and leads to multiple clusters within the spectral feature space (van der Linden et al. 2007). This phenomenon is typical for a category like roof, which comprises a large variety of roofing materials. High inter-class similarity is characteristic for surfaces with high spectral similarity and leads to overlapping class distributions (van der Linden et al. 2007). This phenomenon is typical for

spectrally similar rooftops and paved-areas (e.g., roof bitumen vs. street asphalt), rooftops and pervious surfaces (e.g., clay tiles vs. bare soil), or different vegetation types (e.g., green grass vs. deciduous trees) (compare Figure I-2). On the other hand, urban mapping assessments face the challenge of spectral mixing. This is caused by the heterogeneity of surface types, many of those often with areas smaller than the ground sampling distance of the sensor. As a result, the measured radiation is a composite of the spectral signatures of different materials, which leads to so-called spectrally mixed pixels.

A variety of machine learning algorithms have been applied to remote sensing data, including artificial neural networks (Haykin 1999), decision trees (Quinlan 1986), and support vector machines (Schölkopf and Smola 2002). Machine learning methods optimize their performance by iteratively learning from the data, integrate nonlinear functions, and are non-parametric, i.e., they are not constrained to prior assumptions on data distributions (Waske et al. 2009). They have thus become established and have ever since replaced traditional parametric approaches like the Gaussian maximum likelihood estimator. In the context of urban imaging spectrometry, support vector machines (SVMs) in particular have been demonstrated as powerful analysis techniques. Embedded into a kernel framework and by making use of an optimal adjusted hyperplane (Camps-Valls and Bruzzone 2005), they have proved to be capable of coping with nonlinearities and smooth transition zones between classes, which is required when delineating surface types with high within-class variability and inter-class similarity (van der Linden et al. 2007). The value of the per-pixel support vector machine classifier (SVC) in combination with airborne hyperspectral data for mapping spectrally complex urban surface types has been demonstrated (van der Linden et al. 2007; Waske et al. 2010; Tuia and Camps-Valls 2011). However, assigning a pixel to one thematic category ignores the presence of mixed pixels. This generally produces mapping inaccuracies, even when using high spatial resolution data (Roessner et al. 2001; van der Linden et al. 2007). Due to the dominance of mixed pixels within medium spatial resolution imagery, per-pixel classifiers will likely not be the first choice for urban land cover assessments from future hyperspectral satellite data (Figure I-3).

Quantitative mapping approaches – also referred to as sub-pixel mapping approaches – transform the spectral information of single pixels into physically meaningful quantities of surface fractions and thematically meaningful surface types, and thus account for spectrally mixed pixels. In this context, several sub-pixel mapping techniques have been suggested, including spectral unmixing or regression. Multiple endmember spectral mixture analysis (MESMA; Roberts et al. 1998) is probably the most commonly used unmixing technique

in the urban context (Powell et al. 2007; Franke et al. 2009; Roberts et al. 2012), and relies on the iterative decomposition of mixed pixels into fractional abundances of pure material spectra, i.e., endmembers (EMs). As alternative, the value of different non-parametric and nonlinear machine learning regression algorithms for quantification of urban land cover has been demonstrated (Walton 2008; Esch et al. 2009; Van de Voorde et al. 2009). Methods include the SVM implementation for quantitative prediction, i.e., support vector regression (SVR; Schölkopf and Smola 2002). Compared to MESMA, which only requires a spectral library of EMs, regression approaches rely on appropriate continuous training information, i.e., pairs of spectral signatures and mixing fractions, for function approximation. This information cannot be labeled in the image data itself. A common strategy is to combine pixel spectra with spatially aggregated land cover fractions from a high resolution reference map. However, this approach depends on the availability of such a reference map and on the accurate co-registration of both data sets. Furthermore, the universality of the approach, i.e., the separation of model training – by means of universal spectral databases – from model application and independent validation, is not given. Such a framework is important once the reliability and transferability of regression models to imagery with different spatial resolutions, spatial extents, and acquisition date is explored. Methodological developments toward a universal use of machine learning-based regression approaches are therefore of importance, particularly with regard to the development of operational processing schemes that will be needed once spaceborne hyperspectral data from urban areas become available.

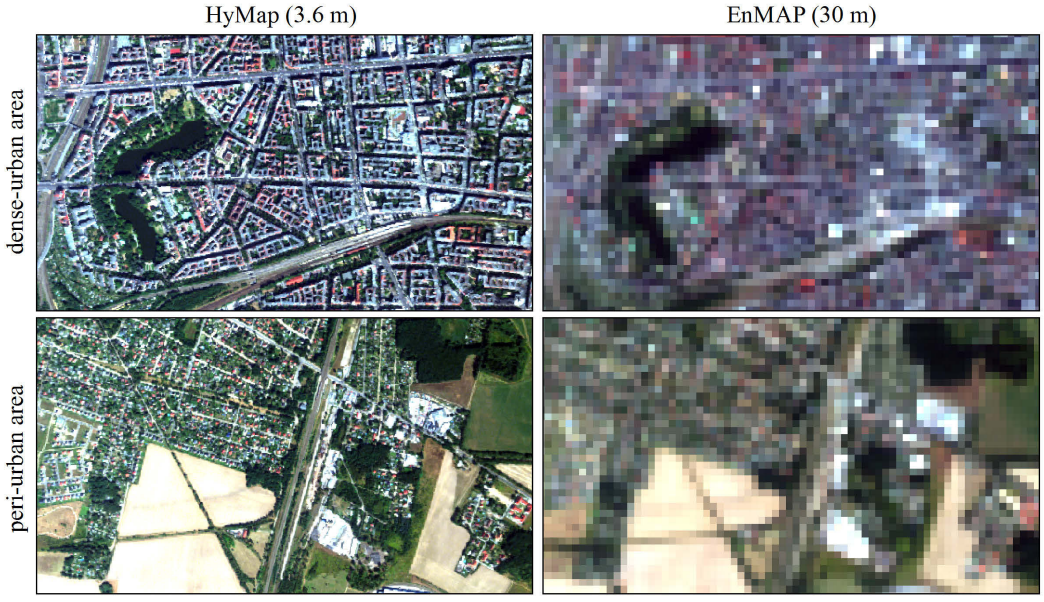


Figure I-3: Extent of spectral mixing within HyMap data at 3.6 m and simulated EnMAP data at 30 m resolution for a dense-urban and a peri-urban area.

2 Conceptual framework

2.1 Motivation

Rapid urbanization and the need to promote the sustainability of cities have created a special challenge for urban remote sensing. The production of maps that capture the diversity of built-up and natural components relevant for urban environmental research or urban management is an essential task of image data processing. Existing global urban map products delineate urban areas by extent, but fail to provide descriptions of different land cover types within the city limits. Future research in urban remote sensing therefore needs to focus on the development of reliable, standardized, and regularly updated mapping products that capture the land cover heterogeneity of human settlements globally.

During the last decade, the value of airborne imaging spectrometer data for mapping detailed urban surface properties has been demonstrated. Data acquired by upcoming hyperspectral satellite missions will create the opportunity to proceed from experimental case studies toward global applications. However, the change in spatial scale from high to coarser spatial resolution hyperspectral data will complicate urban land cover assessments. Investigating benefits of future hyperspectral satellite data over broadly available multispectral satellite data, but also the challenges that will arise from the larger amount of spectrally mixed pixels compared to airborne hyperspectral acquisitions, is of great importance, particularly regarding the potential to derive new urban land cover products.

Operational processing schemes form the basis for deriving standardized map products. With regard to recent developments in image processing techniques, machine learning-based analysis concepts have become increasingly established in remote sensing. In the context of heterogeneous urban environments, advanced quantitative methods to estimate sub-pixel land cover fractions are of particular interest. It is therefore crucial to test and enhance advanced quantitative processing techniques from the field of machine learning with respect to reliability and universality when applied to data of different spatial resolutions, spatial extents, and temporal acquisitions. This may ultimately contribute to the development of an operational processing scheme to derive reliable urban land cover products from future spaceborne imaging spectrometer data.

2.2 Objectives

The scientific focus of this work was to quantify urban land cover by means of machine learning and imaging spectrometer at multiple spatial scales. Experiments considered

innovative methodological developments and opportunities for novel applications. Thus, findings advance current research in urban remote sensing, especially regarding the need for new urban land cover products.

Given the high suitability of machine learning algorithms for effectively exploiting hyperspectral imagery, the first part of this work aimed at developing a universal mapping framework based on SVR for the quantification of urban land cover. The following specific research question was addressed:

Research question I: Does the combination of SVR with synthetically mixed training data allow the reliable quantification of urban land cover from imaging spectrometer data at multiple spatial scales?

To embed SVR into a universal mapping framework, the concept of combining SVR with synthetically mixed training data was introduced. This workflow separates regression model training – based on synthetic data generated from a spectral library – from model application and independent validation, and thus implements the idea of a universal mapping technique. The method's capability to make best use of the latent information content of hyperspectral data was tested by quantifying urban land cover into thematically meaningful, yet spectrally complex urban surface types such as *roof*, *pavement*, *grass*, and *tree*. The method's performance was tested by comparisons to other advanced machine learning regression techniques and MESMA. The method's universality was tested with respect to the adaptability and transferability of regression models when proceeding from high spatial resolution airborne to coarser spatial resolution spaceborne data.

Given the forthcoming availability of hyperspectral satellite data for urban remote sensing, the second part of this work aimed at exploring the potential of the EnMAP mission for urban land cover assessments. The following specific research question was addressed:

Research question II: What type of urban land cover information can be extracted from EnMAP data, and how reliable are the mapping results?

Simulated EnMAP data at 30 m resolution were exploited based on the SVR mapping framework. Emphasis was put on the possibility to map and extend the universal VIS scheme. In addition, benefits and limitations of EnMAP data for urban mapping were assessed by comparison with higher spatial resolution airborne data and with existing multispectral data of similar spatial resolution. The quantification of land cover along the entire urban-rural gradient of Berlin followed a traditional concept for studying ecological

patterns and processes (McDonnell and Pickett 1990). In doing so, a link to the value of EnMAP for urban environmental applications was provided.

2.3 Study area and materials

With approximately 3.3 million people and an area of 892 km², Berlin is the largest city in Germany. The history of Berlin and its hinterland has led to a multifaceted urban-rural environment. During the Second World War, the original fabric of the city was heavily destroyed. Later, during the political division into East- and West-Berlin, a polycentric development with two core areas of differing urban composition characterized the city. After reunification and the establishment of Berlin as the capital of Germany, large scale building activities in central areas and suburbanization through reconnecting the city with its hinterland took place (Sukopp 1990; Schulz 2000). Extensive developments of green spaces have taken place since the 19th century and complement the diversity of the cityscape. Berlin counts more than 2500 public parks, 430,000 street trees, and 930 allotment gardens (SenStadt 2014c).

The study area depicts a subset of Berlin's urban rural-gradient (Figure I-4). More specifically, this region covers an area of approximately 5.5 by 22.8 km and reaches from

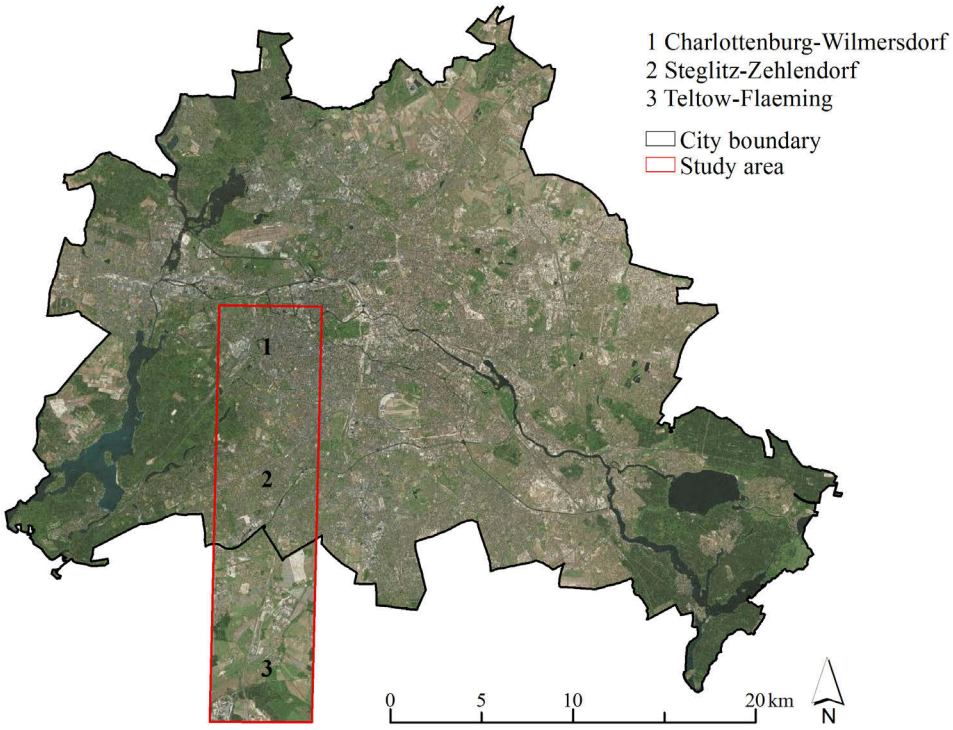


Figure I-4: Study region along Berlin's urban-rural gradient (Background: Digital orthophotos; Source: World Imagery, Esri). The outline of the study region is determined by the spatial extent of the airborne imaging spectrometer data used for experiments.

the inner-city core of the district Charlottenburg-Wilmersdorf (52.51°N, 13.30°E) to the rural area in Brandenburg's district Teltow-Flaeming (52.36° N, 13.31° E). The city area was part of the former West-Berlin and captures a variety of urban structure types of multifunctional use, including a commercial center, residential areas of different densities, industrial zones, sport grounds, green spaces, street trees, and water bodies. The urban hinterland was part of the former German Democratic Republic and is characterized by agricultural land, forest areas, small municipalities, and newly built industrial complexes.

Two HyMap images with 128 spectral bands and 3.6 m and 9 m resolution were acquired during the HyEurope 2009 flight campaign led by the German Aerospace Center. A simulated spaceborne EnMAP scene with 30 m resolution was derived from the 9 m HyMap data. A spectral library consisting of various anthropogenic and natural materials relevant for the study area was derived from the 3.6 m HyMap image. A high resolution reference land cover map for validation was derived using cadastral information, a digital surface model, orthophotos, and field mapping information.

2.4 Structure

This thesis consists of five chapters and one appendix: The Introduction (Chapter I) is followed by three core research chapters (Chapter II-IV), and a Synthesis (Chapter V). The core research chapters (see list below) are stand-alone manuscripts, which were either published or submitted to international peer-reviewed journals. The Synthesis presents a summary of the entire work, draws more general conclusions, and provides applications and future directions for research. The appendix includes a user guide of synthMix-SVR, which is an IDL-based tool that implements the methodological framework developed during this doctoral research.

Chapter II *Okujeni, A., van der Linden, S., Tits, L., Somers, B., & Hostert, P. (2013). Support vector regression and synthetically mixed training data for quantifying urban land cover. Remote Sensing of Environment, 137, 184-197.*

Introduces the concept of combining support vector regression with synthetically mixed training data for quantifying urban land cover using airborne hyperspectral data. A baseline comparison to the traditional Multiple Endmember Spectral Mixture Analysis is provided.

Chapter III *Okujeni, A., van der Linden, S., Jakimow, B., Rabe, A., Verrelst, J., & Hostert, P. (2014). A comparison of advanced regression algorithms for*

quantifying urban land cover. Remote Sensing, 6, 6324-6346.

Compares the utility of five advanced regression techniques combined with synthetically mixed training data for quantifying urban land cover. Airborne hyperspectral data at two different spatial resolutions are used to investigate the possibility to adapt the approach to amplified mixing scenarios.

Chapter IV *Okujeni, A., van der Linden, S., & Hostert, P. (in review)*[*]*. Extending the Vegetation-Impervious-Soil model using simulated EnMAP data and machine learning. Remote Sensing of Environment.*

Illustrates the potential of future EnMAP data for quantifying land cover along the urban-rural gradient. Insights into benefits and limitations are addressed by comparisons to results from higher spatial resolution hyperspectral and coarser spatial resolution multispectral data.

[*] A revised version of this chapter was meanwhile published in: Okujeni, A., van der Linden, S., & Hostert, P. (2015). Extending the vegetation-impervious-soil model using simulated EnMAP data and machine learning. *Remote Sensing of Environment*, 158, 69-80.

Chapter II:
Support vector regression and synthetically mixed training data for quantifying urban land cover

Remote Sensing of Environment, Volume 137, October 2013, Pages 184–197

Akpona Okujeni, Sebastian van der Linden, Laurent Tits, Ben Somers and Patrick Hostert

© 2013 Elsevier Inc. All rights reserved.
DOI: 10.1016/j.rse.2013.06.007
Received 16 October 2012; Revised 24 April 2013; Accepted 9 June 2013; Available online 19 July 2013.

Abstract

Exploiting imaging spectrometer data with machine learning algorithms has been demonstrated to be an excellent choice for mapping ecologically meaningful land cover categories in spectrally complex urban environments. However, the potential of kernel-based regression techniques for quantitatively analyzing urban composition has not yet been fully explored. To a great extent, this can be explained by difficulties in deriving quantitative training information that reliably represents pairs of spectral signatures with associated land cover fractions needed for empirical modeling. In this paper we present an approach to circumvent this limitation by combining support vector regression (SVR) with synthetically mixed training data to map sub-pixel fractions of single urban land cover categories of interest. This approach was tested on Hyperspectral Mapper (HyMap) data acquired over Berlin, Germany. Fraction estimates were validated with extensive manual mappings and compared to fractions derived from Multiple Endmember Spectral Mixture Analysis (MESMA). Our regression results demonstrate that the sets of multiple mixtures yielded high accuracies for quantitative estimates for four spectrally complex urban land cover types, i.e., fractions of impervious rooftops and pavements, as well as grass- and tree-covered areas. Despite the extrapolation uncertainty of SVR, which resulted in fraction values below 0% and above 100%, physically meaningful model outputs were reported for a clear majority of pixels, and visual inspection underpinned the quality of produced fraction maps. Statistical accuracy assessment with detailed reference information for 92 urban blocks showed linear relations with R^2 values of 0.86, 0.58, 0.81 and 0.85 for the four categories, respectively. Mean absolute errors (MAE) ranged from 6.4 to 12.8% and block-wise sums of the four individually modeled category fractions were always around 100%. Results of MESMA followed similar trends, but with slightly lower accuracies. Our findings demonstrate that the combination of SVR and synthetically mixed training data enable the use of empirical regression for sub-pixel mapping. Thus, the strengths of kernel-based approaches for quantifying urban land cover from imaging spectrometer data can be well utilized. Remaining uncertainties and limitations were related to the known phenomena of spectral similarity or ambiguity of urban materials, the spectral deficiencies in shaded areas, or the dependency on comprehensive and representative spectral libraries. Therefore, the suggested workflow constitutes a new flexible and extendable universal modeling approach to map land cover fractions.

1 Introduction

Research on the structure and functioning of cities is of great importance, particularly considering the rates of current and projected global urbanization (Alberti 2005; Grimm et al. 2008; Pickett et al. 2011). Detailed descriptions of urban surface properties, i.e., beyond the broad differentiation of “urban land” and “non-urban environments”, are crucial for developing a more thorough understanding of the integrated human-natural ecosystem (Cadenasso et al. 2007). The spatial distribution, abundances and conditions of different impervious, pervious and vegetation cover types, for example, directly influence processes related to urban climate (Chudnovsky et al. 2004; Gluch et al. 2006), hydrology (Arnold and Gibbons 1996; Pauleit and Duhme 2000) or biodiversity (Blair 1996; McKinney 2002).

Data from municipal authorities provide detailed information on urban surface characteristics at the scale of urban blocks, a preferred spatial unit for urban management purposes or urban environmental studies (Pauleit and Duhme 2000; Heiden et al. 2012). Data collections are, however, often based on cost- and time-intensive field surveys and are subject to irregular updates. In this context, remote sensing constitutes a supplemental technique to derive information on the physical composition of urban areas (Jensen and Cowen 1999; Miller and Small 2003; Maktav et al. 2005).

Data acquired by various multispectral spaceborne sensors, which range from very high to moderate spatial resolution and high acquisition frequency, have been applied for mapping urban land cover (Rashed et al. 2003; Thomas et al. 2003; Powell et al. 2007; Myint et al. 2011). Many studies followed the V-I-S model by Ridd (1995), a framework used to universally characterize three major urban surface components, i.e., “vegetation”, “impervious surfaces” and “soil”. However, the coarse spectral resolution of multispectral data often limits the accurate distinction of different impervious, pervious and vegetation surface cover types (Herold et al. 2003; Small and Lu 2006; Myint et al. 2011). Therefore, Small (2004) and Small and Lu (2006) proposed the Substrate Vegetation Dark surface (SVD) model, which characterizes the urban environment by biophysical surface reflectance properties rather than by land cover types. Nevertheless, linking spectral with thematic information remains desirable, especially when high resolution spectral information is available. Other studies overcame some of the limitations of multispectral data by analyzing high-resolution images with object-based strategies (Shackelford and

Davis 2003; Thomas et al. 2003; Myint et al. 2011). Still others use complex multi-sensor and multi-temporal approaches that extend the spectral feature space or take advantage of information beyond the spectral domain (Griffiths et al. 2010; Taubenböck et al. 2012). Even when employing such fusion approaches we need to better understand the opportunities and limitations of high resolution spectral data.

The increasing availability of imaging spectrometer data from urban areas has allowed for a more detailed mapping of thematic surface properties on a purely spectral basis. Several studies have demonstrated the great potential of continuous spectral information for fine-scale classifications, including various man-made materials and vegetation species (Herold et al. 2003; Franke et al. 2009). Other studies have included rather broad-scale classifications when considering ecologically meaningful sub-categories, e.g., impervious rooftops and pavements, pervious soils, and grass- and tree-covered areas (Roessner et al. 2001; van der Linden and Hostert 2009). Thus far such investigations have been more experimental in nature because currently available high quality data sets are almost exclusively constrained to airborne acquisitions with limited spatial coverage and temporal frequency. This is likely to change with the advent of new hyperspectral satellite missions, e.g., the German Environmental Mapping and Analysis Program (EnMAP; Stuffler et al. 2007; Stuffler et al. 2009) or the Hyperspectral Infrared Imager (HyspIRI; NRC 2007; Roberts et al. 2012). Such systems will provide imaging spectrometer data of a large spatial extent on a timely and frequent basis, which will open up new opportunities for urban remote sensing applications (Heldens et al. 2011). Ground sampling distances (GSDs) in the range of 30 to 60 m will, however, complicate studies, especially in spatially heterogeneous environments. Robust analysis approaches with universal applicability and best possible transferability on data of coarser spatial resolution will be needed to best utilize imagery acquired by future spaceborne imaging spectrometers.

Despite this promising development, using imaging spectrometer data for urban applications remains a challenge. The high number of materials, their varying conditions, as well as the surfaces' anisotropic reflectance behavior lead to vast spectral diversity. As a consequence, different spectral phenomena such as high within-class variability of individual materials, complex multi-modal class compositions, and high inter-class similarity or ambiguity complicate urban land cover assessments (Herold et al. 2004; Herold et al. 2006; Schiefer et al. 2006; Heiden et al. 2007). This complexity is further aggravated by the high amount of spectrally mixed pixels typical for urban remote sensing data. The extent of spectral mixing, i.e., the number of mixed pixels and the abundance of

different materials contributing to the mixed signal, strongly depends on the fine-scale spatial patterns of different urban objects, and on the GSD of the respective sensor (Small and Lu 2006).

Spectrally mixed signatures produce inaccuracies in per-pixel classification approaches, which assign each pixel to a discrete urban land cover category (Small 2001; Powell et al. 2007). This applies even to investigations on airborne or high resolution spaceborne remote sensing data of 4 m GSD (Small 2003; van der Linden et al. 2007). Hence, quantitative mapping of urban land cover components constitutes an alternative concept that becomes particularly relevant once coarser resolution spaceborne imaging spectrometer data becomes available. In general, quantitative approaches account for sub-pixel mixing by transforming reflectance measurements of individual pixels into both physically meaningful quantities of surface fractions and thematically meaningful land cover types. Multiple Endmember Spectral Mixture Analysis (MESMA; Roberts et al. 1998) is probably the most commonly used technique to systematically decompose mixed pixels into fractional abundances of distinct spectrally pure endmembers (EMs). Within-class spectral variability is accounted for by using an extensive iterative procedure with multiple linear mixture models (Somers et al. 2011). In the context of imaging spectrometry of urban areas, MESMA was successfully used to quantify thematically detailed land cover types or to derive material-oriented fraction maps (Franke et al. 2009; Roberts et al. 2012).

Support vector machines (SVMs) have received increasing attention in the remote sensing community; SVMs are supervised, non-parametric statistical learning techniques designed to solve classification and regression problems (Vapnik 1995; Schölkopf and Smola 2002). A recent review of implementations and applications of SVMs was provided by Mountrakis et al. (2011). As applied to remote sensing data, SVMs have shown great capabilities to model complex multi-modal, nonlinear data distributions in high-dimensional spectral feature spaces (Huang et al. 2002; Melgani and Bruzzone 2004; Pal and Mather 2006). Thus far, the advantage of kernel-based support vector classification (SVC) for a per-pixel-based mapping of spectrally complex urban land cover categories from imaging spectrometer data has been demonstrated (van der Linden et al. 2007; Waske et al. 2010; Tuia and Camps-Valls 2011). However, the potential of kernel-based support vector regression (SVR) as a quantitative technique for analyzing the complex hyperdimensional urban feature space with high spectral and spatial heterogeneity has not yet been fully explored. Given a set of adequate training samples, SVR allows us to estimate a continuous output variable, such as sub-pixel fractions of a single land cover

category. To date, most studies have used SVR for predicting biophysical/chemical plant parameters by relating spectral information to in situ samples (Camps-Valls et al. 2006; Tuia et al. 2011; Verrelst et al. 2012). Few studies have adopted SVR to quantify fractions of broad urban categories in 30 m multispectral satellite imagery using reference information from very high resolution land cover maps (Walton 2008; Esch et al. 2009). This latter approach fully relies on the availability of very accurately co-registered, very high resolution reference. Often such data – and especially accuracy – is not given when high resolution imagery from spatially heterogeneous environments is used as input, e.g., for airborne line scanner data from urban areas. Moreover, reliable sub-pixel fractions cannot be labeled in the data itself or mapped in the field. These difficulties in finding reliable quantitative training information explain the lack of studies that use regression techniques for sub-pixel mapping. To the best of our knowledge, no investigations have been carried out where SVR was used together with imaging spectrometer data to derive sub-pixel fraction maps of spectrally complex urban land cover categories.

In this paper, we propose the combination of support vector regression with synthetically mixed training data to quantify urban land cover. The idea behind our approach was inspired by Foody and Mathur (2006), who suggested generating class-dominated spectral mixtures using spectral averaging to improve SVC training. Based on imaging spectrometer data and a corresponding spectral library, we consecutively derived fraction maps with SVR for single urban land cover categories of interest, so-called target categories. For each target category we generated a synthetically mixed training data set from the spectral library, which was subsequently used for SVR model training. This training set consists of pure original spectra and multiple-binary mixed spectra representing a range of mixing fractions between 0 and 100% of the target category against all remaining background cover types. To generate a universal model that accounts for spectral variability, the spectral variations within each land cover category, e.g., varying illumination and shade, or differences in material composition and condition, were included and learned from appropriate training. We thus fulfilled the need for quantitative training information required for empirical modeling, and we overcame the mentioned difficulties of using regression techniques for sub-pixel mapping. Subsequently, each SVR model was applied to the image data to derive the desired fraction map per target category.

The objective of this study was to quantify spectrally complex and ecologically meaningful urban land cover categories using a combination of SVR and synthetically mixed training data. We thus explore whether (machine learning-based) regression techniques can be used

for sub-pixel mapping with training data derived from commonly available sources such as spectral libraries. We used airborne Hyperspectral Mapper (HyMap) data acquired over Berlin, Germany, to consecutively estimate fractions of two different impervious categories, i.e., rooftops (*roof*) and paved surfaces (*pavement*), including roads, sidewalks and parking lots. We further differentiated two vegetation categories, i.e., *grass* and *tree*. Our objective was tested by comparing fraction estimates at pixel and urban block scale to maps from very high resolution reference data. To explore the quality of results compared to established approaches we evaluated our maps to results derived with MESMA.

2 Data

2.1 Imaging spectrometer data and pre-processing

Image data was acquired by the HyMap sensor over Berlin on 20th August 2009 at approximately 11:45 local time. HyMap covers the spectral region between 440 and 2500 nm, with 128 bands and bandwidths between 10 and 20 nm. The sensor's field of view (FOV) is 60°, and the instantaneous field of view (IFOV) is 2.5 mr along track and 2.0 mr across track (Cocks et al. 1998). The aircraft was flown from south to north at an altitude of 2005 m (AGL), resulting in raw pixels of 5 × 4 m at nadir. Given this setup, an area of 2.3 by 22.8 km was covered, representing a subset of Berlin's urban-rural gradient (Figure II-1). A great variety of urban structure types were captured, including a commercial center, residential areas of varying densities, green spaces, water bodies, sport grounds, industrial areas, as well as agricultural areas and forest patches.

Pre-processing of the HyMap image encompassed system correction (Cocks et al. 1998) and atmospheric correction (Richter and Schläpfer 2002) by HyVista Corporation and the German Aerospace Center (DLR). Data were delivered in reflectance values with a reduced number of 126 bands. We performed a parametric geocoding (Schläpfer and Richter 2002) using a 3.6 m resolution digital surface model (DSM; acquired in January 2008 and resampled from 0.5 m). This cell size also constitutes the final spatial resolution of the ortho-rectified HyMap image. The slightly higher output resolution than that of the raw image was chosen to better preserve a higher portion of the spectral information during geometric processing (Schläpfer and Richter 2002). Based on 15 independent control points from digital orthophotos (DOPs; acquired in April 2009 with 0.1 m resolution), we calculated a spatial accuracy (RMSE) of 2.91 and 3.08 m in easting and northing.

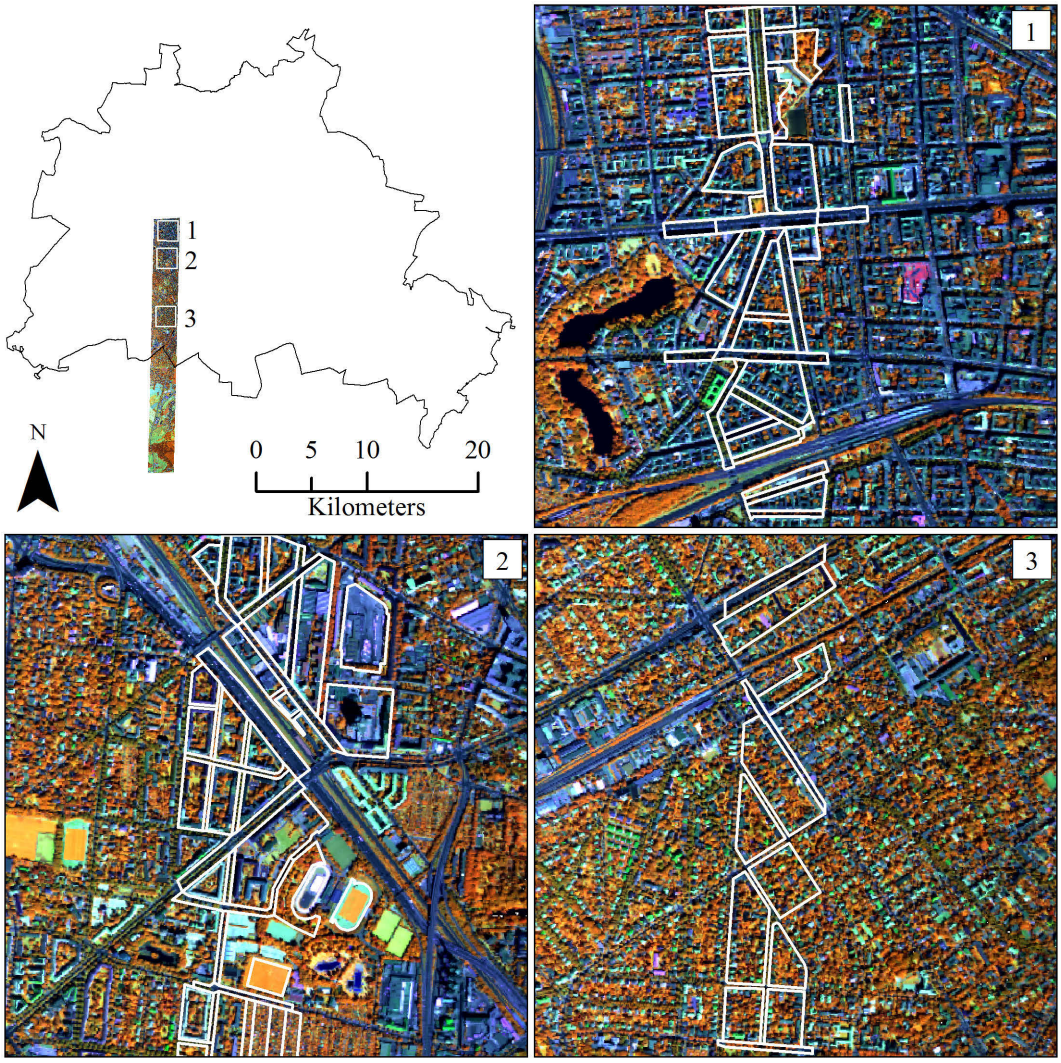


Figure II-1: HyMap footprint covering a subset of Berlin's urban gradient (R = 833 nm, G = 1652 nm, B = 632 nm) and validation areas 1-3 (white polygons indicate urban blocks used for validation).

2.2 Image spectral library

A comprehensive spectral library was developed from the HyMap image. We included major characteristic man-made and natural surface materials relevant for the study area. The library was structured according to a slightly modified hierarchical categorization scheme for urban areas as proposed by Roessner et al. (2001) and Heiden et al. (2007). Levels I and II represent rather broad land cover categories such as impervious, pervious or vegetation. *Roof*, *pavement*, *grass* and *tree* correspond to Level III and represent the more thematically detailed urban land cover categories of interest in this work. Level IV forms the spectral basis for the analysis (Table II-1).

The initial collection of around 300 pure material spectra (mean values of manually delineated spectrally homogeneous regions of interest in the HyMap image) was based on

Table II-1: Hierarchically structured image spectral library.

Level III	Level IV	No. spectra
Roof	Red clay tile	4 (a)
	Red cement tile	3 (a)
	Bitumen	5 (b)
	Brown roof tile	1
	Brown roof shingle	1
	White roof material (polyethylene)	1
	White roof material (unknown)	1
	Zinc roof material	1
Pavement	Asphalt	4 (a, b)
	Concrete	2 (b)
Grass	Grass (intensively manicured)	2 (a)
	Grass (extensively manicured)	1
	Grass (dry)	2 (a)
Tree	Deciduous tree	7 (a)
Other	Tartan (sports ground)	1
	Railtrack (concrete sleepers)	1
	Railtrack (wooden sleepers)	1
	Sand (playground)	1
	Soil	1
	Water	1

Multiple spectra per material considered to account for spectral variability due to variations in (a) illumination and shading effects, and (b) material condition or composition.

expert knowledge and auxiliary information, including field mapping data, DOPs and Google Street View. We considered the spectral variability of materials, e.g., variations in brightness caused by varying illumination and shading effects, or variations due to differences in material condition and composition. In a second step, we iteratively developed the final image spectral library used for processing. To create a subset of spectra from the initial collection that best describes the spectral diversity of the study region, we manually selected a representative standard spectrum (spectrum close to the mean within the collection of multiple material spectra) for materials with frequent spatial occurrence. We adopted this subset of spectra to derive fraction maps using SVR and synthetically mixed training data. After visual inspection and statistical accuracy assessment of results, we added spectra to account for missing materials. Spectra were also added to account for illumination and shading effects or variations in condition and composition. The final image spectral library consists of 41 material spectra, of which 17 were assigned to the *roof* category, 6 to the *pavement* category, 5 to the *grass* category and 7 to the *tree* category. Six spectra were assigned to a so-called *other* category. This additional category, with mainly non-vegetated pervious surfaces of marginal spatial occurrence, did not constitute a category of interest for this analysis.

2.3 Reference data

Most analyses of urban environmental indicators by municipal authorities in Berlin are performed on a block scale, and related maps are presented in the Urban and Environmental Information System (UEIS; SenStadt 2013a). These urban blocks, therefore, formed the spatial unit to statistically validate the results in three regions of various urban structure types (Figure II-1). Validation area 1 represents a high-density commercial and housing zone (high-density urban area), area 2 represents a medium-density residential and industrial zone (medium-density urban area), and area 3 represents a low-density residential zone with mainly detached and semi-detached housing patterns and private gardens (low-density urban area). The effect of urban surface geometry, e.g., building displacement at large view-angles, constitutes a significant drawback for any wide-swath sensor over urban areas. When compared to reference data (e.g., cadastral information), the three-dimensional geometry produces inaccuracies related not to the spectral information content, but rather to unavoidable geometric distortions. We therefore only selected blocks within a ±10° off-nadir region to minimize effects of the urban 3D-geometry on comparisons between HyMap and reference data. By limiting our evaluation to that area of the scene where almost entirely spectral effects (and not geometric) influence our accuracy measures, we were able to determine the spectral effectiveness of the SVR approach. The effects of urban 3D-geometry on wide-swath sensor data and the related performance limits of any algorithm in urban environments is beyond the scope of this work and exhaustively described in van der Linden and Hostert (2009). Reference fractions of Level III land cover categories of interest (see Table II-1) were calculated for each urban block. The spatial extent of buildings was extracted from Berlin's digital cadastre (SenStadt 2013b). Vegetation types and paved impervious grounds were manually delineated at 1:500 map scale based on the available DSM and DOPs. Whereas the building data set was regarded as spatiotemporal-invariant, the temporal shift of the DSM and DOPs to the Hymap image may have influenced validation to some extent. The final reference data set encompassed a total number of 92 urban blocks with areas between 2,530 and 65,950 m², of which there were 45 building blocks of different urban structure type, 37 street polygons including wide main streets and small side roads, and 10 green spaces consisting of parks, allotment gardens or sport grounds.

3 Methods and data analysis

The complete methodological framework of this study is illustrated in Figure II-2. Imaging spectrometer data, a corresponding spectral library and independent reference information formed the data base for processing. Initially, library spectra were used to generate synthetic mixtures, which served, along with their mixing fractions, as input for SVR model training. Cover fraction estimates were then derived by applying the SVR model to the image data. The entire procedure was carried out separately for each of the four land cover categories of interest. Finally, accuracy was assessed with an independent reference data set.

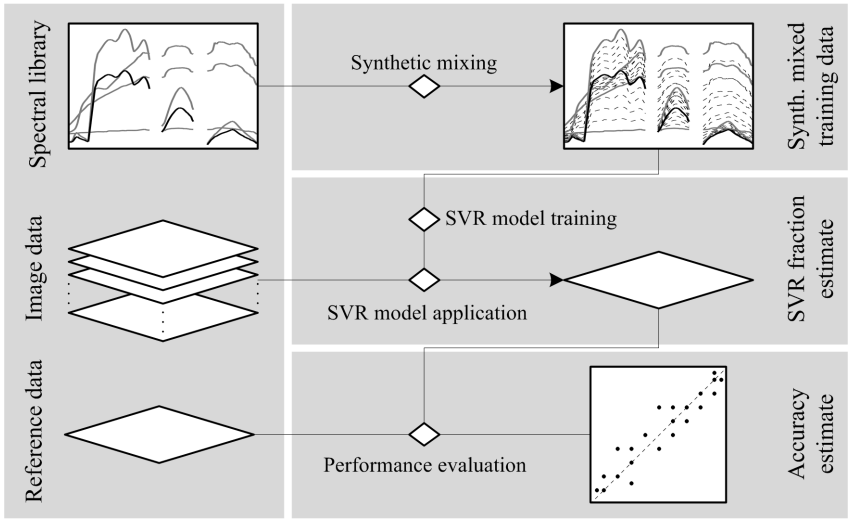


Figure II-2: Workflow for quantifying urban land cover with SVR and synthetically mixed training data.

3.1 Generation of synthetically mixed training data

The generation of synthetically mixed training data from a spectral library is illustrated in Figure II-3. For each Level III urban land cover category of interest, a quantitative training set was generated as follows:

Step 1: The image spectral library was split into two sections, (i) a target category and (ii) a background category, which included all remaining categories.

Step 2: Synthetically mixed spectra between each pure spectrum of the target category (i.e., 100% mixing fraction), and each pure spectrum of the background category (i.e., 0% mixing fraction of the current target category) were calculated. In line with many unmixing approaches (e.g., Tompkins et al. 1997; Roberts et al. 1998), we hereby assumed linear mixing systematics for simplification. While

constructing the synthetically mixed training data, decisions on the mixing complexity and the mixing interval had to be made.

- The mixing complexity refers to the number of materials from the target and background category that contribute to a mixed pixel. In general this is not known in advance, and thus requires a flexible strategy to account for both simple and more complex mixtures. Binary mixtures between each pure spectrum of the target category and each pure spectrum of the background category constitute the simplest complexity. This complexity can be arbitrarily increased by allowing intra- and inter-categorical mixtures with more than two materials, e.g., three or four to calculate ternary or quaternary mixtures. In this study, we approximated the mixing complexity inherent in the HyMap image by overlaying a 3.6 m grid on the reference data set. Our calculations indicated that less than 10% of the pixels are composed of more than two land cover types. Although this analysis could not be carried out on a material level, we used this approximation to investigate the use of binary mixtures for deriving reliable fraction estimates from the 3.6 m HyMap image.
- The mixing interval refers to the number of intermediate mixtures within the given fraction range between 0 and 100% when mixing the target category spectra against the background category spectra. This interval can be defined globally or for each land cover type separately, either by a constant step size or by arbitrarily selected percentages. To keep the computing time low, the definition of the mixing interval should further balance the trade-off between the necessary and avoidable number of intermediate mixing steps, while at the same time considering the accuracy of fraction estimates. In this study we trained SVR models with global step sizes of 50%, 25%, 20%, 10% and 5%. Our results showed that accuracies increased with decreasing increments. For step sizes below 20%, no considerable performance improvements were observed. However, a clear increase in processing time was observed. Therefore, we ultimately selected the global step size of 20%.

Step 3: Finally, all pure original and binary-mixed spectra were combined in a single spectral library. This multiple set of spectra, together with the associated mixing fractions between 0 and 100% of the target category, constituted the training data for SVR modeling.

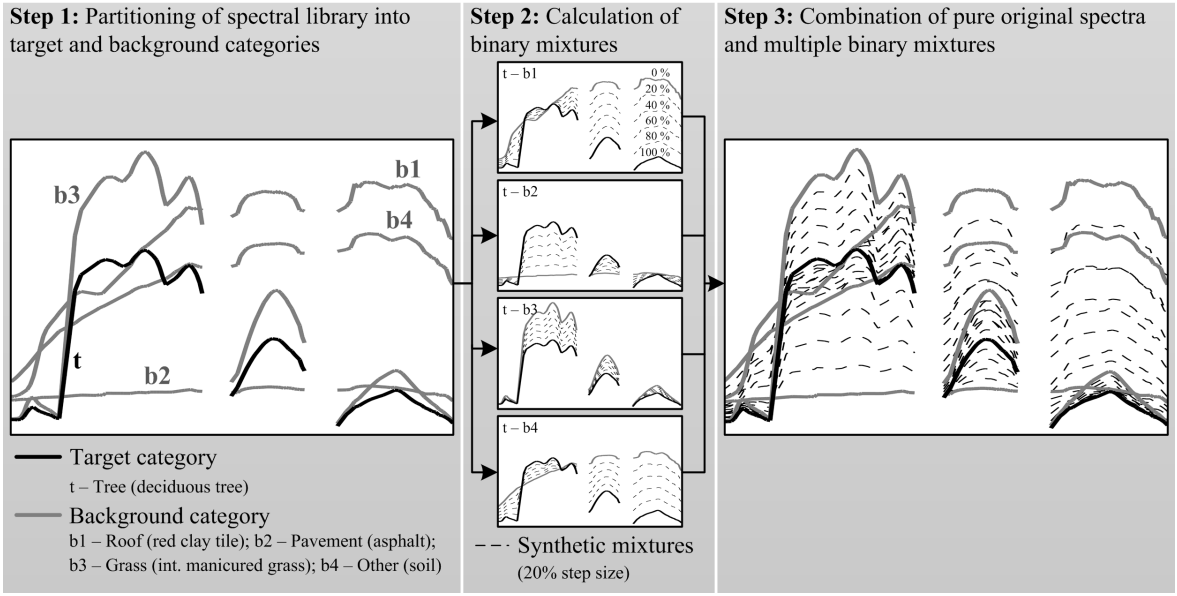


Figure II-3: Generation of synthetically mixed training data.

3.2 Support vector regression

The main idea behind SVR is to estimate the linear dependency between pairs of n-dimensional input vectors and 1-dimensional target variables by fitting an optimal approximating hyperplane to a set of training samples. The hyperplane is defined by a linear function that is found based on structural risk minimization. That is, the linear model is optimal when it minimizes a cost function considering a maximized margin that encloses the samples, and a minimized error of approximation. Approximation is controlled by an ε-insensitive loss function that directly influences the width of the margin. Samples within the margin are not penalized. The error of approximation, i.e., samples located outside of the margin, is measured using slack variables. The influence of samples outside the margin is controlled by a regularization parameter *C*. Integrated into a kernel framework, SVR is capable of capturing nonlinear data distributions. By using a Gaussian kernel function, the data is implicitly mapped into a higher dimensional feature space, wherein the new data distribution enables a better fitting of a linear function. Hence, kernel-based SVR turns out to be a nonlinear regression in the original input space (Vapnik 1995; Smola and Schölkopf 2004; Brereton and Lloyd 2010).

The SVR analysis was carried out with imageSVM (Rabe et al. 2010a), which is an open source tool for support vector machine applications with remote sensing data. Delivered as part of the EnMAP-Box (Rabe et al. 2012), imageSVM uses LIBSVM (Chang and Lin 2011) during the training of support vector models. We followed a two-step training approach: initially, for each target category the synthetically mixed data set was used as

input to train an SVR model. Optimal parameter sets for γ, which determines the width of the Gaussian kernel, as well as for C and ε, were determined by a grid search in conjunction with an internal 3-fold cross-validation using the RMSE as a measure of performance. Thus, the effectiveness of SVR to find a single solution, i.e., one global model, for the complex functional relation between the multiple set of spectra and associated mixing fractions was exploited. Subsequently, each SVR model was applied to the HyMap image, which resulted in one fraction image per target category. The model output range covers continuous values between 0 and 100%, i.e., partial interpolations of 20% training intervals, as well as values below 0% or greater than 100%, i.e., extrapolated model outputs.

3.3 Multiple endmember spectral mixture analysis

A more physically-based approach for estimating sub-pixel cover fractions is MESMA (Roberts et al. 1998). MESMA is directly driven by the physically explicit mixture model that describes the mixed pixel signal as a linear combination of the spectral properties of the constituent components (or EMs) weighted by their relative cover in the pixel. Once the EMs and their spectral signatures are known, inversion techniques are used to solve for the fractions with minimal additional error in the model equations, thus obtaining fully constrained cover fractions. As opposed to more traditional unmixing techniques, MESMA allows the EMs to vary on a per-pixel basis. Many mixture models are iteratively calculated for each pixel in an image, thereby evaluating all possible EM combinations from a spectral library and calculating the fit between the measured and modeled mixed signal. The model with the lowest RMSE is assigned to the pixel. Often, the mixing complexity is handled by calculating simple two-EM models and more complex three- and four-EM models per pixel. The final model is selected based on a given criterion, e.g., a model of higher complexity is only selected if the error decrease related to increased complexity exceeds a threshold (Powell et al. 2007; Franke et al. 2009; Roberts et al. 2012).

The MESMA approach aims to achieve a comprehensive estimation of land cover fractions for all urban categories considered in the analysis. The 41 pure library spectra related to Level III of the hierarchical categorization scheme (Table II-1) were used as endmembers for processing. As less than 10% of the pixels are composed of more than two land cover types (compare Section 3.1), all possible combinations between two material endmembers were evaluated. Consequently, we do not account for higher order mixing complexities. We

do, however, guarantee a fair methodological comparison between both approaches in terms of mixing complexity considered in the analysis. In the initial setup, shadow was not accounted for, as shaded materials were included in the library (Asner and Lobell 2000). This resulted in a total number of 623 two-EM models (two material EMs). In the second setup, we included a spectrally flat photometric shade spectrum as additional EM, with subsequent shade normalization (Powell et al. 2007; Roberts et al. 2012). This resulted in a total number of 623 three-EM models (two material EMs, one shade EM). For each setup, results of the mixture model with lowest modeling error per pixel were retained, with the corresponding estimated cover fractions for the land cover categories considered in the analysis, i.e., *roof*, *pavement*, *grass*, *tree* and *other*.

3.4 Validation of fraction maps

Pixel-wise accuracy assessment of fraction estimates from empirical modeling is not straightforward. Similar to the dilemma related to acquiring quantitative training data, inaccurate co-registration between mapping results and higher resolution reference information will lead to biased accuracy statistics, especially in such a spatially heterogeneous urban environment. Spatial units larger than a pixel, e.g., a cluster of pixels or polygons, can mitigate this influence (Stehman and Wickham 2011). This strategy was used, for example, by Powell et al. (2007) and Roberts et al. (2012) to statistically validate MESMA-based urban land cover fraction maps. In addition to a pixel-wise evaluation, i.e., visual inspection of fraction maps and a consistency check of fraction values due to the extrapolation uncertainty of SVR, we used urban blocks from the UEIS to statistically assess the accuracy of our results.

4 Results

4.1 Performance of SVR at pixel scale

The consistency of pixel values was used to evaluate the reliability of fraction estimates derived by SVR and its ability to generalize. We considered the percentages of physically meaningful (0–100%), negative (<0%) and super-positive (>100%) fractions for all urban blocks and for urban blocks stratified by structure types (Table II-2). For all urban blocks, around 60% of *roof* or *pavement* pixels show physically meaningful fractions between 0 and 100% (Table II-2a). The percentage of negative fractions for both categories is clearly higher than the percentage of super-positive fractions. With over 80% of pixels, *grass* and

Table II-2: Percentages of physically meaningful, negative and super-positive pixel fractions for a) all urban blocks, and b) urban blocks stratified by structure types.

	Category	0-100%	<0%	>100%
a) All				
	Roof	62.9	27.5	9.6
	Pavement	58.8	24.9	16.3
	Grass	80.6	19.0	0.4
	Tree	81.5	15.2	3.4
b) By urban structure type				
High- density	*Roof*	61.3	21.3	17.4
	Pavement	54.5	30.8	14.8
	Grass	67.2	32.7	0.0
	Tree	73.3	24.6	2.1
Medium- density	*Roof*	64.0	23.3	12.7
	Pavement	59.1	25.7	15.2
	Grass	76.2	23.4	0.3
	Tree	79.8	18.3	1.9
Low- density	*Roof*	66.1	28.1	5.8
	Pavement	64.1	21.8	14.1
	Grass	91.7	8.2	0.2
	Tree	91.5	6.0	2.4
Street	*Roof*	56.5	39.0	4.4
	Pavement	56.6	19.8	23.7
	Grass	85.3	14.7	0.0
	Tree	80.8	11.6	7.6
Green space	*Roof*	75.3	23.7	1.1
	Pavement	63.1	30.8	6.1
	Grass	93.0	3.3	3.7
	Tree	86.7	10.9	2.4

tree show a higher number of physically meaningful fractions. Stratified results further illustrate that percentages vary between urban structure types (Table II-2b). For all land cover categories, the number of physically meaningful fractions increases along with a decreasing density of urban block. At more than 90%, the highest increase can be observed for *grass* and *tree* within urban blocks from low-density urban areas. With approximately 56% of *roof* and *pavement* pixels, street polygons show the lowest numbers of physically meaningful fractions.

An analysis of the spatial distribution of pixels with physically unrealistic fractions values reveals that for all urban land cover types, most of the pixels with super-positive fractions are associated with pure surface coverage of the respective target category (Figure II-4). Vice versa, most of the pixels with negative fraction values are linked to surfaces with target fractions of 0%. Minor deviations from this common trend can be observed for *roof*

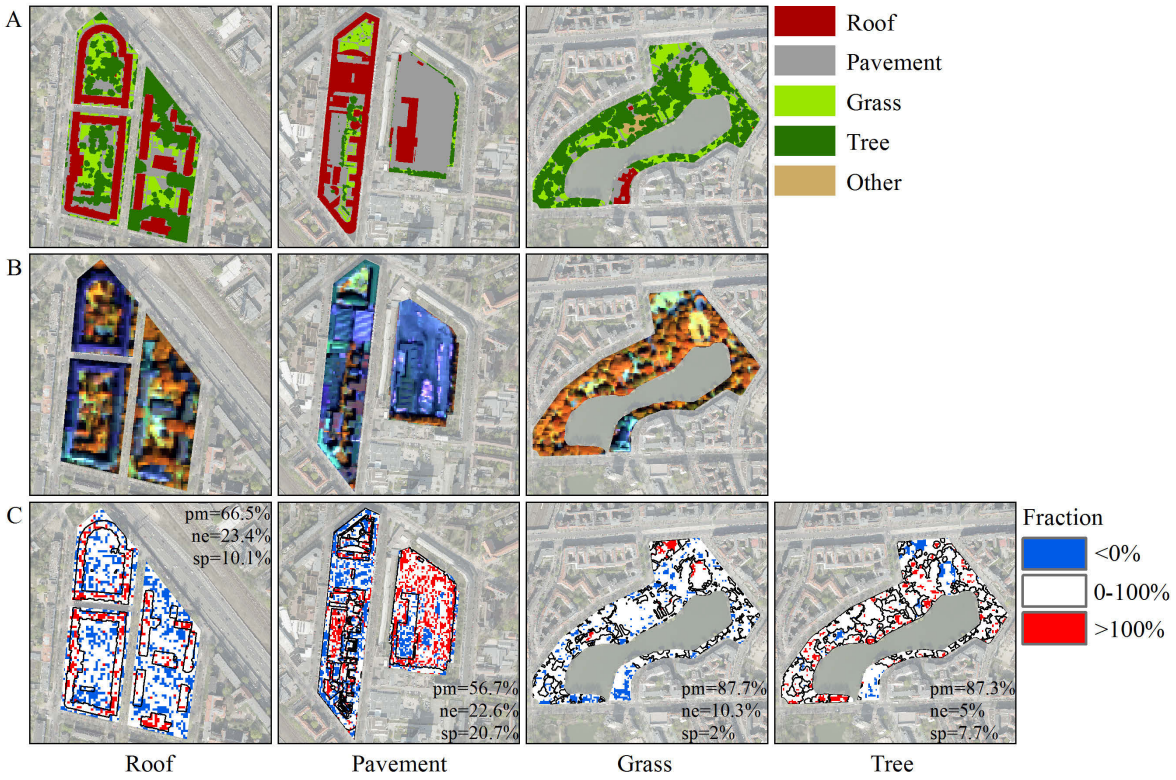


Figure II-4: Subsets showing the spatial distribution of pixels with physically meaningful and unrealistic fractions values for different urban blocks. (A) Reference data, (B) HyMap image (R = 833 nm, G = 1652 nm, B = 632 nm) and, (C) aggregated fractions modeled by SVR. The percentage of physically meaningful (pm), negative (ne) and super-positive (sp) fractions is reported in each image chip.

and *pavement*, e.g., rooftops or parts of a paved bus parking area that are misleadingly mapped with negative fractions.

Given this rather meaningful location of outliers, we produced stretched fraction maps with values between 0 and 100%, where negative and super-positive fractions were set to 0 and 100%, respectively. Visual inspection of five selected subsets reveals that spatial patterns of the urban land cover categories in the reference data set are reproduced well by SVR (Figure II-5). For example, *roof* is differentiated from the background cover types with high precision. Minor confusion occurs in the case of paved streets and backyards or shaded areas, i.e., areas covered by cast shadows of buildings and trees or shaded components of trees. In a few cases, particularly for dark roofing materials, *roof* fractions are underrated. The *pavement* estimates are accurate for streets or backyards. Simultaneously, however, a weak quantitative differentiation, especially from dark background cover types is observed. This leads to highly overrated *pavement* fractions for some rooftops and, above all for shaded grass- or tree-covered areas. These effects are particularly strong for highly fragmented low-density urban areas and green spaces, both of which have complex illumination and shadowing effects within vegetation stands. The

quantitative differentiation of *grass* and *tree* from background cover types by SVR is highly accurate. Different types of grass-covered surfaces, e.g., lawns consisting of dry or extensively manicured grass, are correctly mapped as having full *grass* coverage of 100%. The precise distinction between *grass* and *tree* can be observed along streets, i.e., where street trees dominate the scenery, and in the example of green space. Confusion between the two vegetation types is relatively low, and mostly appears when tree crowns are highly illuminated.

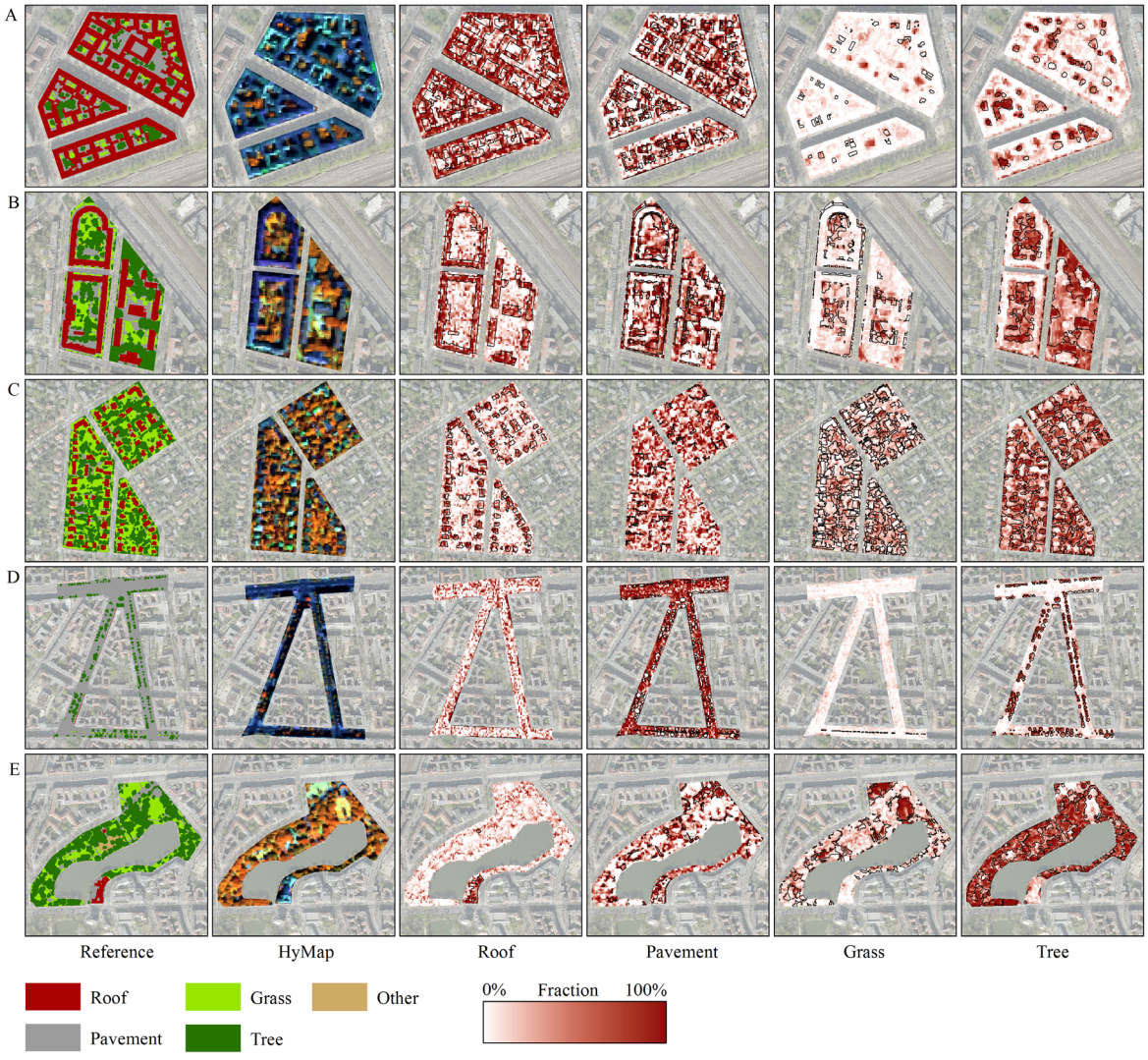


Figure II-5: Subsets showing urban blocks of different structure types from the reference data, HyMap image (R = 833 nm, G = 1652 nm, B = 632 nm) and fraction maps modeled with SVR. (A) High-density-, (B) medium-density-, (C) low-density urban area, (D) street polygons and, (E) green space. Fraction images are stretched between 0 and 100% (values <0% set to 0%, values >100% set to 100%).

4.2 Performance of SVR and MESMA at urban block scale

Urban land cover fraction estimates modeled by SVR and MESMA are compared to reference information at the urban block scale. We used the original SVR outputs, i.e., unstretched fraction values, to calculate average block estimates. For a total of 92 urban blocks, *roof*, *grass* and *tree* estimates by SVR show linear relationships close to the 1:1 line, with high accuracies as indicated by MAE values below 9% and R^2 values over 0.8 (Figure II-6). Estimates of *pavement* are less accurate with increased scattering around the 1:1 line, a higher MAE value of 12.8% and a lower R^2 value of 0.58. For most urban blocks, SVR produces physically meaningful fraction values between 0 and 100%. Only five urban blocks show negative values of up to -5%.

For all urban land cover types, accuracy measures of MESMA carried out in the setup with two-EM models (two material EMs) follow similar trends but with slightly reduced accuracies when compared to SVR (Figure II-6). The largest deviations from SVR results are observed for *pavement*, where the MAE increases by 7.2%. In terms of R^2, a quite large decrease of 0.2 is observed for *grass* estimates. As mentioned in Section 3.3, MESMA was additionally carried out in a setup with three-EM models (two material EMs, one shade EM) with subsequent brightness normalization. Results are very similar (scatterplot not shown), with both exhibiting slightly increased and reduced accuracies (*roof*: MAE = 11.7%, R^2 = 0.72; *pavement*: MAE = 20.1%, R^2 = 0.61; *grass*: MAE = 9.5%, R^2 = 0.59; *tree*: MAE = 6.3%, R^2 = 0.75).

The scatterplots (Figure II-6) reveal regions of agreement and disagreement between both approaches. The *roof* estimates of most urban blocks from medium- and high-density urban areas are highly underrated by MESMA. For both methods, street polygons and green spaces are misleadingly mapped as having *roof* proportions of up to 20%. However, *pavement* estimates by SVR and MESMA show opposite tendencies. For building blocks from low-density urban areas and green spaces, *pavement* proportions derived by SVR are highly overestimated. In contrast, MESMA highly underrates intermediate to high *pavement* fractions, especially for street polygons. The *grass* and *tree* estimates of both approaches follow similar patterns. The reduced accuracies of MESMA compared to SVR for both categories can be attributed to small differences. The *grass* fractions, particularly of building blocks from low-density urban areas, are slightly underestimated by SVR, and to a greater extent by MESMA. The effect of misleadingly mapped *grass* fractions within street polygons is higher for MESMA. Further, *tree* fractions within street polygons are

underrated by MESMA. For both approaches, the largest underestimates of *grass* and overestimates of *tree* fractions are observed for four green spaces with highly fragmented allotment gardens. Also, in this case deviations from the 1:1 line are larger for MESMA than for SVR.

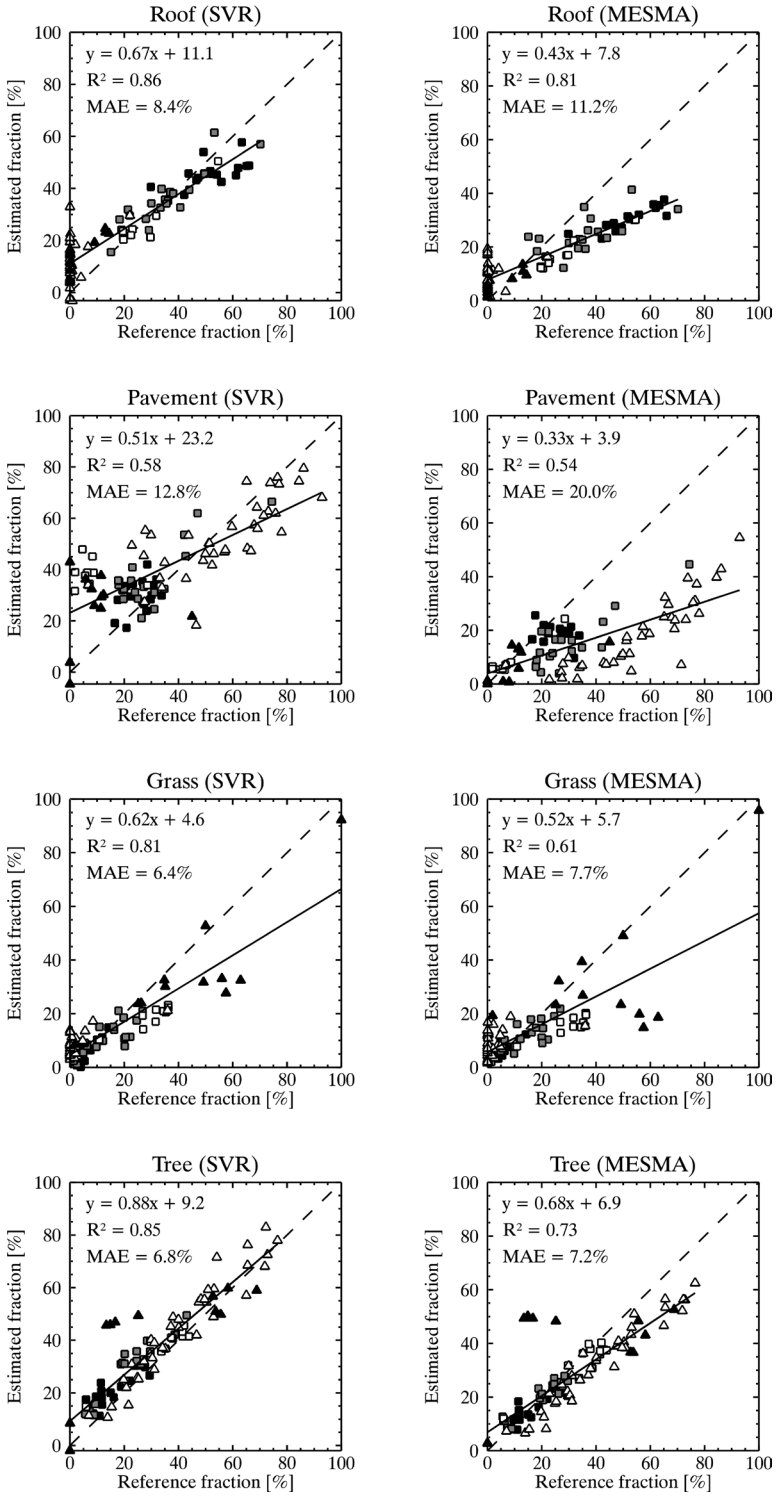


Figure II-6: Scatterplots of *roof*, *pavement*, *grass* and *tree* fractions for 92 urban blocks from the reference data set compared to modeled SVR and MESMA (two-EM models) fractions. Square symbols = building blocks (black = high-density-, gray = medium-density-, white = low-density urban area); triangle symbols = green spaces (black), street polygons (white).

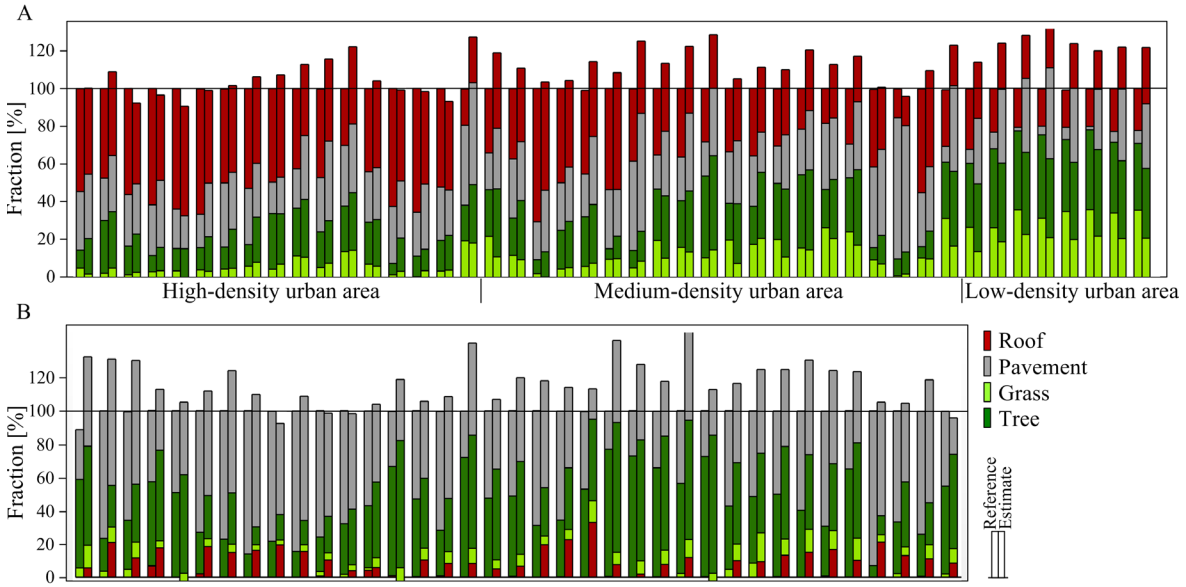


Figure II-7: Combined *roof*, *pavement*, *grass* and *tree* fraction estimates for (A) 45 building blocks and (B) 37 street polygons.

The reliability of the proposed SVR approach for quantitatively mapping urban land cover at the urban block scale was further evaluated by investigating whether estimated fractions that were consecutively derived by four different SVR models sum to unity. For 45 building blocks, block-wise sums of *roof*, *pavement*, *grass* and *tree* estimates produce reasonable fractions with a mean value of 111.4% and a standard deviation of 10.9% (Figure II-7a). Block-wise sums of street polygons are slightly less accurate with a mean value of 116.7% and a standard deviation of 13.1% (Figure II-7b). The reasons for super-positive sums accord with previously observed results. For building blocks, physically unrealistic fractions clearly above 100% mainly occur when *pavement* is highly overestimated. This effect is particularly strong in low-density urban areas. For street polygons, block-wise sums clearly above 100% are mainly caused by misleadingly estimated *roof* and *grass* fractions, which are either inexistent or negligibly small in the reference data set.

5 Discussion

5.1 Urban land cover fractions estimates by SVR at pixel scale

Accurate percentages of physically meaningful fraction values at the pixel scale highlight the predictive power of the SVR models and prove the capability of the sets of pure and synthetically mixed spectra to account for large parts of the spectral diversity within the

HyMap image. The fact that spectrally complex, i.e., multi-modal, urban land cover categories such as *roof* and *pavement*, as well as spectrally similar classes, e.g., *tree* and *grass*, can be quantitatively separated shows the strength of SVR to extract almost all spectral information. Although statistical assessment cannot be performed at the pixel scale, the visual inspection of maps (see Figure II-5) strengthens the proposed SVR approach.

Nevertheless, two major types of errors exist during SVR modeling. First, unconsidered spectra of potential surface materials or material variability can lead to improper extrapolation, i.e., the estimation beyond the known training space. Second, the spectral ambiguity of training data can lead to interpolation uncertainties when fitting the hyperplane. Both errors lead to reduced mapping performances, either by inaccurate land cover fraction estimates in the physically meaningful range between 0 and 100%, or by unrealistic fraction value estimates, i.e., negative or super-positive values (see Table II-2).

Accordingly, a high number of physically unrealistic fraction values are found where spectral variability of man-made materials is high, e.g., for the *roof* category or for dense urban areas where complex illumination and shadowing effects amplify spectral variations. Similarly, negative and super-positive *roof* and *pavement* fraction values often occur in areas where spectral ambiguity between the target and background category is high. This is particularly the case for urban blocks that include large paved areas, e.g., parking lots (see Figure II-4) or streets (see Table II-2). In contrast, *grass* and *tree* predominantly show physically meaningful fraction values and only very few super-positive fractions. High accuracies for vegetated surfaces underline the approach's capability to quantitatively describe diverse categories with often subtle spectral differences.

Despite the unfavorable existence of outliers, the comparison of reference data and the pixel-wise spatial distribution of physically unrealistic fraction values indicate that this problem can be compensated for in a post-processing step: by setting negative and super-positive fraction values to 0% and 100%, respectively, physically meaningful cover fractions of the target category are achieved (see Figure II-5).

5.2 Urban land cover fraction estimates by SVR and MESMA at urban block scale

Accurate maps of urban surface properties, i.e., fractional abundances of various types of impervious and vegetated surface cover types, serve as an important data basis for proper urban planning or ecological studies (Pauleit and Duhme 2000; Cadenasso et al. 2007; Heiden et al. 2012). Aggregated land cover information at the block scale constitutes a

suitable spatial unit for such purposes (Pauleit and Duhme 2000; Heiden et al. 2012). We statistically evaluated sub-pixel fraction estimates by SVR at the urban block level. The quality of results was further evaluated by means of a comparison to MESMA. By (i) using the same image spectral library, and (ii) considering the same mixing complexity, we created identical starting points for a fair methodological comparison.

The high accuracies achieved at the urban block scale underscore the capability of SVR, along with the synthetically mixed training data, to quantitatively map urban land cover (see Figure II-6). For a total number of 92 urban blocks we differentiated four spectrally complex and similar target categories with high precision. Reduced accuracies of *pavement* estimates generally indicate the critical mapping of paved surface types like roads, sidewalks or parking lots. Effective discrimination between surfaces from the target category and from the remaining background cover types is illustrated by the reproduced spatial patterns of the SVR-based fraction maps (see Figure II-5). The high performance of the proposed SVR approach is particularly noteworthy for the spectrally complex *roof* category, which includes a diversity of roofing materials at different illumination conditions. The same applies for the approach's capability to handle both vegetation categories, *grass* and *tree*. For example, grass surfaces with differences in physiological condition were mapped as desired with maximal fractions. Further, trees along street corridors or within green spaces were in large part unambiguously quantified from background cover types. Inaccurate mappings of vegetation types were low and occasionally occurred when brightly illuminated tree crowns were confused with photosynthetically active grass. This applies, for example, to individual sun-exposed trees along streets or in backyards, parking lots and the like. Inaccurate mappings of both vegetation types for the allotment gardens can be, to a great extent, explained by uncertainties in the reference data, where the interactive mapping of highly fragmented patches of grass and trees appears critical. Adapting the synthetically mixed training data to more detailed vegetation characteristics, i.e., distinguishing green vegetation from non-photosynthetic vegetation, or including vegetation species, could be easily incorporated but was beyond the scope of this work.

Similar to previous studies (Franke et al. 2009; Roberts et al. 2012), MESMA is demonstrated to be a powerful technique for the quantitative mapping of urban land cover. This is underscored by the accurate results, particularly for *roof*, *grass* and *tree* estimates (see Figure II-6). Likewise, the quantitative description of the *pavement* category appears critical. Performances are very similar for two different setups, i.e., two-EM models (two

material EMs) and three-EM models (two material EMs, one shade EM) with subsequent shade normalization.

When directly comparing the results of both approaches, SVR achieves higher accuracies at the urban block scale (see Figure II-6). However, considering the reference data, both negative and super-positive fraction values at the pixel scale (see Figure II-4) have a positive balancing effect on the calculation of urban block mean values. This must be taken into account when interpreting the SVR-based fraction estimates at the urban block scale. The SVR and MESMA mappings show common sources of errors, which are generally linked to the limitations of urban land cover assessments on a spectral basis, even when using imaging spectrometer data. Misleadingly mapped *roof* fractions along street corridors can be attributed to the spectral similarity or ambiguity of many impervious surface types, particularly to the well-known confusion between darker bituminous roofing materials and street asphalt (Herold et al. 2004; van der Linden et al. 2007; Franke et al. 2009), or to the high numbers of cars, which can be spectrally more similar to roofing materials (van der Linden et al. 2007). Similarly, mapped *roof* fractions within green spaces can be attributed to the confusion between spectra of dark roofing materials and spectra of shaded parts of vegetation stands. The lower accuracies of *pavement* can be ascribed to two major sources of uncertainty. First, paved surfaces contain a higher portion of spectrally ambiguous materials, leading to more spectral confusion. In contrast, the higher number of distinct spectra in the *roof*, *grass* and *tree* categories leads to unambiguous fraction estimates. Second, the high spectral similarity between materials from the *pavement* category, shaded surfaces and water bodies plays an important role, but with fundamental differences between both methods. Using SVR and synthetically mixed training data, shadow-covered grass surfaces or shaded components of trees are misleadingly mapped as having high *pavement* fractions. This effect is particularly visible in urban blocks with high proportions of vegetation, i.e., in low density urban areas or green spaces, and can be attributed to the spectral confusion between the characteristic dark flat spectral signature of asphalt and the low reflectance of shaded areas. In contrast, MESMA underestimates intermediate and high *pavement* fractions, particularly along streets. This is due to the almost zero reflectance water spectrum within the *other* category, which is universally used by MESMA to model non-water pixels with low flat reflectance signatures, e.g., asphalt. Hence, underrated *pavement* fractions are erroneously compensated by high *other* fractions. Including a spectrally flat photometric shade spectrum in the three-EM model (two material EMs, one shade EM) with subsequent shade normalization led to slightly

improved *pavement* results, but did not compensate for the observed underestimation along streets. This is due to the similarity of the water and the photometric shade signature. Hence, the use of an additional shade EM is only effective when excluding the water spectrum from the library.

Given these observations, several similarities can be identified for both approaches. In general, the quality of the SVR and MESMA mappings depends on the availability of appropriate training information during modeling. The comprehensiveness of the library spectra ensures the reliable estimation of land cover fractions from the image. Further, common sources of uncertainty for both SVR- and MESMA-based mappings are often related to spectral ambiguities. This shortcoming cannot be overcome by any analysis technique when using spectral information alone. Moreover, both approaches assume that spectrally mixed pixels can be approximated by linear mixing systematics. The high achieved accuracies support this simplification, as well as the omission of possible nonlinear mixing effects occurring from multiple-scattering between objects in the sensors' field of view (Borel and Gerstl 1994; Somers et al. 2009a). Finally, both methods can be flexibly adapted to the spectral complexity of the analyzed image data. This is if of great importance with respect to the coarser spatial resolution of future spaceborne imaging spectrometer data, where mixtures beyond two urban land cover categories will prevail in most pixels. In terms of SVR, synthetically mixed training data can be created both for simple binary mixtures and for more complex ones, e.g., ternary or quaternary mixtures. Similarly, MESMA can exploit EM models of different complexities. In addition, expert knowledge can be used to optionally exclude specific mixtures that do not occur in nature.

Despite these similarities, SVR is conceptually different from MESMA. These differences have implications on the implementation and performance of the respective technique, and thus on the ultimate selection of the appropriate method for quantifying urban land cover. First, modeling urban land cover with SVR requires quantitative training information, which is derived by synthetically mixed training data from the spectral library. This enables empirical modeling with regression, as alternative reference information is difficult to collect. MESMA uses pure spectra as EM for estimating fractions by model inversion. Accordingly, no further interaction by the user is necessary. Second, with SVR we generally aim to estimate land cover fractions of specific target categories. Depending on the study purpose, this can be either one individual urban land cover category such as "urban tree", or several subsequently mapped land cover categories to comprehensively describe the physical composition of urban areas. The quality of the proposed SVR

approach for comprehensive mapping is demonstrated by the feasibility of combining the fraction maps produced by four different SVR models without imposing constraints (see Figure II-7). One general advantage of SVR – the ability to model the complex functional relations between the set of multiple binary-mixtures and associated mixing fractions in a one-step approach – is exploited. A single global SVR model per target category is trained to map urban land cover, while at the same time accounting for spectral variability. In contrast, MESMA aims to achieve a comprehensive estimation of cover fractions for all urban categories considered in the analysis. Spectral variability is accounted for by iteratively calculating all possible two-EM mixture model combinations on a per-pixel basis. The model with the best fit (i.e., RMSE between observed and modeled spectra) is selected. As all land cover categories in the analysis are considered, MESMA generates an additional fraction map visualizing the fractional abundances of the *other* category, a land cover category of minor interest in this work. Third, accuracies of modeling results depend on the capability of the respective technique to exploit the spectral information given during training. Thus, the general strengths of SVR to cope with spectral complexity and to effectively exploit subtle spectral differences between otherwise spectrally similar materials by individually weighting spectral bands are utilized. This may to a great extent explain the more accurate SVR-based fraction estimates, particularly for *roof*, *grass* and *tree* estimates. In contrast, MESMA equally weights all wavebands, and thus is subject to the risk of underemphasizing these possibly decisive spectral features (Somers et al. 2009b). The difference in performance between both methods generally indicates the strengths of SVR to optimally exploit the spectral information provided by the training input. Therefore, the excellent suitability of SVMs for mapping complex urban land cover types is again demonstrated.

6 Conclusion

In this paper we demonstrate the potential of the machine learning-based SVR to be used for the quantitative mapping of four spectrally complex and ecologically meaningful urban land cover types on a purely spectral basis. Imaging spectrometer data was used to estimate fractions of impervious rooftops and pavements, as well as grass- and tree-covered areas along the urban gradient of Berlin, Germany. We propose the integration of synthetically mixed training data into the training stage of the analysis. This poses a straightforward and repeatable strategy to overcome the dilemma related to the difficult acquisition of

quantitative training information that is needed for sub-pixel mapping with empirical regression techniques.

Constructing synthetically mixed data for SVR model training only requires a spectral library that represents the spectral diversity of the study area. Using an image spectral library allows easy associations to be made with spectral features of the image. Universal spectral databases, e.g., consisting of laboratory or field spectra of man-made urban materials and vegetation species, can be used alternatively, providing that radiometric consistency between library spectra and the image is given. This becomes important when the number of pure spectra present in the image is low, e.g., as expected for future spaceborne imaging spectrometer data of urban areas.

The objective of this study, i.e., to quantify the four spectrally complex and ecologically meaningful urban land cover categories using the proposed SVR approach, was tested by comparing fraction estimates at the pixel and the urban block scale to manual field mappings, and to fraction maps derived by MESMA. Although conceptually different, both methods demonstrate their great potential to quantify the land cover of spectrally heterogeneous urban environments. The proposed SVR approach aims to quantify a single target land cover category in a universal global model that accounts for spectral variability. The subsequent analysis of several target categories allows for a reliable comprehensive mapping of urban land cover. This is different from the traditional MESMA method, which uses pure endmembers in an extensive iterative procedure with multiple linear mixture models for the comprehensive mapping of all land cover categories considered in the analysis. The high accuracies achieved in this study demonstrate that the combination of SVR and synthetically mixed training data can be recommended as an alternative quantitative mapping method, particularly when users are only interested in mapping specific land cover categories of interest. Moreover, improved fraction estimates at the urban block scale by SVR indicate the general strengths of machine-learning-based techniques to effectively cope with the spectral complexity inherent in imaging spectrometer data from urban areas.

Most of the observed uncertainties can be related to the known phenomena of spectral ambiguity or similarity of urban materials and the spectral deficiencies in shaded areas. Whereas spectral ambiguities cannot be overcome by any analysis technique when using spectral information alone, strategies to cope with spectral similarities or shadow effects must be developed, particularly when analyzing impervious surface cover types. Future

work should therefore focus on data transformation (e.g.; Wu 2004) or spectral feature extraction techniques (e.g.; Herold et al. 2004; Heiden et al. 2007) prior to SVR processing.

The presented method was tested in a spectrally and spatially challenging urban environment (van der Linden and Hostert 2009). The robustness and accuracy of the approach can therefore be expected to be of similar or better quality in more homogeneous areas. Given the powerful separation of two vegetations types, the quantitative mapping of natural ecosystems appears especially worthwhile. By separating the model training process from the model application and validation, the general idea of a universal quantitative mapping approach is given. This is important for future studies, when accuracy, reliability and transferability of SVR models to image data with different spatial extents, spatial resolution and varying acquisition dates must be tested.

Acknowledgements

This research is funded by the German Research Foundation (DFG) under project no. HO 2568/2-2 and partly funded by the German Aerospace Centre (DLR) and the Federal Ministry of Economics and Technology (BMWi) as part of the EnMAP Core Science Team activities (FKZ 50EE0949). The authors would like to thank the two anonymous reviewers for their valuable comments. The authors are grateful to M. Cierpinski for generating the reference data, A. Rabe, B. Jakimow and S. Suess for the support during SVR analysis and helpful comments. The contribution of the students from Humboldt-Universität zu Berlin, who assisted with the field survey, is greatly appreciated. The HyMap data set was acquired and pre-processed by the DLR in the framework of the HyEurope 2009 campaign. HyEurope 2009 was funded by the German Federal Ministry of Education and Research (BMBF). We thank the Berlin Senate Department for Urban Development for providing data from the Urban and Environmental Information System (UEIS) and the DSM.

Chapter III:
A comparison of advanced regression algorithms for quantifying urban land cover

Remote Sensing, Volume 6, Issue 7, July 2014, Pages 6324–6346

Akpona Okujeni, Sebastian van der Linden, Benjamin Jakimow, Andreas Rabe, Jochem Verrelst and Patrick Hostert

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).
DOI: 10.3390/rs6076324
Received 31 March 2014; Revised 30 June 2014; Accepted 2 July 2014; Published 7 July 2014.

Abstract

Quantitative methods for mapping sub-pixel land cover fractions are gaining increasing attention, particularly with regard to upcoming hyperspectral satellite missions. We evaluated five advanced regression algorithms combined with synthetically mixed training data for quantifying urban land cover from HyMap data at 3.6 and 9 m spatial resolution. Methods included support vector regression (SVR), kernel ridge regression (KRR), artificial neural networks (NN), random forest regression (RFR) and partial least squares regression (PLSR). Our experiments demonstrate that both kernel methods SVR and KRR yield high accuracies for mapping complex urban surface types, i.e., rooftops, pavements, grass- and tree-covered areas. SVR and KRR models proved to be stable with regard to the spatial and spectral differences between both images and effectively utilized the higher complexity of the synthetic training mixtures for improving estimates for coarser resolution data. Observed deficiencies mainly relate to known problems arising from spectral similarities or shadowing. The remaining regressors either revealed erratic (NN) or limited (RFR and PLSR) performances when comprehensively mapping urban land cover. Our findings suggest that the combination of kernel-based regression methods, such as SVR and KRR, with synthetically mixed training data is well suited for quantifying urban land cover from imaging spectrometer data at multiple scales.

1 Introduction

Urban areas represent highly challenging environments for remote sensing data analysis. Compared to natural environments, urban areas are very heterogeneous featuring a high spectral diversity of various anthropogenic and natural materials (Herold et al. 2004; Heiden et al. 2007) and a spatially complex, three-dimensional surface geometry (Small 2003; van der Linden and Hostert 2009). To separate meaningful impervious and vegetation sub-categories on a purely spectral basis (Roessner et al. 2001; Franke et al. 2009), imaging spectrometer data is often the only adequate source. Results in Herold et al. (2003) and Gamba and Dell'Acqua (2006) illustrate the surplus of hyperspectral data for mapping different urban surface types when compared to results from multispectral imagery. In addition to the spectral requirements, special attention has to be paid to the selection of appropriate analysis techniques that can cope with spectral similarities, spectrally complex class compositions (van der Linden et al. 2007) and the mixed pixel problem typical for urban remote sensing data (Small 2003; Powell et al. 2007).

In terms of per-pixel classifications from hyperspectral imagery, in which each pixel is assigned to a discrete land cover type, powerful approaches from the field of machine learning have been widely studied (Melgani and Bruzzone 2004; Camps-Valls and Bruzzone 2005; Pal and Mather 2006; Clark and Roberts 2012; Colgan et al. 2012; Im et al. 2012a). Machine learning techniques effectively deal with high dimensional data and learn the relationship between variables by fitting flexible, non-parametric and nonlinear models without a priori assumptions on data distributions (Schölkopf and Smola 2002; Camps-Valls and Bruzzone 2009). These properties are important when generating urban land cover maps that differentiate roofs (i.e., a spectrally complex multi-modal class) from paved areas and soils (i.e., classes with high spectral similarities). The same applies to the mapping of different vegetation types such as grass- and tree-covered areas. Particularly, the kernel-based support vector machine classifier was identified as a robust technique for producing accurate maps of spectrally complex and similar urban land cover categories from airborne imaging spectrometer data (van der Linden et al. 2007; Tuia and Camps-Valls 2011).

Sub-pixel mapping of urban land cover constitutes an alternative concept. Unlike per-pixel classifiers, sub-pixel methods account for the mixed pixel problem by decomposing the signature of a pixel into physically meaningful quantities of surface fractions and

thematically meaningful land cover types. This is of importance considering that even in higher spatial resolution airborne hyperspectral data, the presence of spectral mixtures was identified as one major source of error when employing per-pixel classifiers (Roessner et al. 2001; van der Linden et al. 2007). Multiple endmember spectral mixture analysis (MESMA; Roberts et al. 1998) is probably the most commonly used method to quantify fractional abundances of spectrally pure endmembers (EMs) of a given spectral library. Spectral variability is accounted for by iteratively calculating multiple linear mixture models using all possible EM combinations on a per-pixel basis. The mixing complexity is accounted for by allowing the number of EMs to vary (e.g., 2-EM models, 3-EM models, etc.). The model with the best fit between measured and modeled signal is ultimately selected. MESMA has been successfully adopted for quantifying urban surface properties using imaging spectrometer data (Franke et al. 2009; Roberts et al. 2012).

Regression techniques provide continuous outputs at pixel scale. They are thus also well suited for estimating sub-pixel land cover fractions and serve as alternative to spectral unmixing. Empirical regression modeling relies on the availability of continuous training information, i.e., pairs of spectral signatures and related land cover fractions. In general, this information cannot be labeled in the data itself or mapped in the field. A suitable strategy to derive appropriate training data is to combine image spectra with spatially aggregated land cover information from high resolution reference data. This approach was frequently adopted for mapping broad urban land cover categories, e.g., impervious or vegetation cover, from coarser resolution multispectral satellite imagery. Methods range from multiple regression techniques (Bauer et al. 2008; Van de Voorde et al. 2011) to advanced machine learning techniques such as artificial neural networks (Pu et al. 2008; Van de Voorde et al. 2009), regression trees (Yang et al. 2003; Yuan et al. 2008; Im et al. 2012b) or support vector regression (Walton 2008; Esch et al. 2009).

However, little attention has been paid to the use of regression approaches for sub-pixel mapping by means of imaging spectrometer data. Although advanced regression techniques incorporate benefits of machine learning (e.g., non-parametric and nonlinear modeling) necessary for mapping spectrally complex land cover types, they were primarily utilized for estimating biophysical/chemical plant parameters (Bacour et al. 2006; Camps-Valls et al. 2006; Verrelst et al. 2012; Cernicharo et al. 2013; Yu et al. 2013). To some extent, this shortcoming relates to difficulties in finding reliable training information. The use of spatially aggregated fractions relies on accurately co-registered image and reference data sets and small spatial shifts may lead to biased fraction representations. While the

effect of misregistration is less pronounced when adopting coarser spatial resolution spaceborne imagery, this constitutes a problem when exploiting higher resolution airborne imaging spectrometer data from heterogeneous environments.

In a recent study (Okujeni et al. 2013), we introduced the strategy of combining support vector regression with synthetically mixed training data for mapping sub-pixel fractions of single urban land cover categories, so-called target categories. Synthetically mixed training data are sets of pure library spectra, multiple mixed spectra as well as related mixing proportions, which are used as input for subsequent regression modeling. While generating synthetically mixed training data, spectral variability is accounted for by utilizing the spectral variations of materials included in the library. Mixtures can be flexibly calculated at different complexity levels, e.g., simple mixtures between two spectra and more complex ones between multiple spectra, to account for the mixing complexity inherent in the image data. By using such a strategy, we solved the problem of deriving quantitative training information needed for empirical modeling. Moreover, we were able to utilize the strength of the machine learning-based support vector regression by means of synthetic training data. Combining regression techniques with synthetic training data is conceptually different to MESMA. The complex functional relation between multiple linear mixtures and related mixing fractions is solved in a single global model per target category. The subsequent analysis of several target categories allows for comprehensively mapping urban land cover. In contrast, MESMA uses an iterative procedure with multiple linear mixture models for mapping all land cover categories considered in the analysis. A baseline comparison to MESMA in Okujeni et al. (2013) demonstrated that the combination of support vector regression and synthetically mixed training data improved the accuracy of sub-pixel fraction estimates in a complex urban environment and, hence, was recommended as an alternative quantitative mapping method.

So far, currently available high quality imaging spectrometer data sets are almost exclusively constrained to airborne acquisitions with limited spatial coverage and temporal frequency. This constitutes a significant limitation for more applied urban land cover assessments by means of hyperspectral information. With the advent of new spaceborne imaging spectrometers, e.g., the Environmental Mapping and Analysis Program (EnMAP; Stuffler et al. 2009) or the Hyperspectral Infrared Imager (HyspIRI; Roberts et al. 2012), hyperspectral imagery will become broadly available. Acquisitions from space, however, come along with coarser spatial resolution compared to the currently dominating airborne data sets. New quantitative analysis approaches that best cope with coarser spatial

resolution and even more abundant mixed pixels are therefore required to utilize these data for unprecedented possibilities. In this context, the training of advanced regression models from field or auxiliary information with subsequent upscaling from airborne to spaceborne imagery constitutes a flexible strategy for universal quantitative modeling. Hence, both (i) research on the potential of advanced machine learning regression techniques for exploiting urban imaging spectrometer data and (ii) investigations on the influence of decreasing spatial resolution on the quality of urban mapping appear worthwhile.

In this work, we further investigated the potential of synthetically generated training data for exploiting urban imaging spectrometer data. We aim at providing a more generic evaluation of the concept of synthetically mixed training data to be used for urban mapping. We utilized five different advanced regression techniques, including kernel-based support vector regression (SVR; Schölkopf and Smola 2002), kernel ridge regression (KRR; Hastie et al. 2009), artificial neural networks (NN; Haykin 1999), random forest regression (RFR; Breiman 2001) and partial least squares regression (PLSR; Wold et al. 2001). Analyses were carried out on hyperspectral data at two different spatial resolutions acquired over Berlin, Germany. We considered four typical urban land cover categories of interest, namely the two impervious surface types *roof* and *pavement* as well as the two vegetation types *grass* and *tree*. The quantification of urbanized land into such sub-categories plays an important role for urban environmental research or urban planning (Pauleit and Duhme 2000; Cadenasso et al. 2007). Specifically, our study addresses the following research questions:

(1) Which regression techniques effectively utilize synthetically mixed training data for producing accurate urban land cover maps?

(2) How do coarser resolution images influence the performance of regression models?

(3) Do more complex synthetic mixtures allow for coping with potential deficiencies of coarser resolution imagery?

2 Study area and materials

2.1 Study area

The study region covers a subset of the urban–rural gradient of Berlin, Germany (Figure III-1). Along this transect we selected three validation areas of different urban structure types, which include manifold man-made materials and vegetation types typical for the

city. The “high-density urban area” represents an inner-city commercial and residential zone with dense building block structures. The “medium-density urban area” includes residential zones with less dense perimeter block or row house structures as well as industrial zones with large, paved areas. The “low-density urban area” represents a suburban residential zone with mainly detached and semi-detached housing patterns with associated private gardens. All three validation areas include green spaces and street corridors of different sizes. Given this makeup, a representative spectrally and spatially challenging heterogeneous urban environment was considered for our experiments.

2.2 Image data

We used imaging spectrometer data acquired by the HyMap sensor during the HyEurope 2009 Campaign of the German Aerospace Centre (DLR) on 20 August 2009. The aircraft was flown over the study area twice, with identical nadir position but at different altitudes of 2005 and 4750 m. This resulted in two independently acquired scenes with spatial resolutions of 3.6 m (HyMap01) and 9 m (HyMap02) after image ortho-rectification

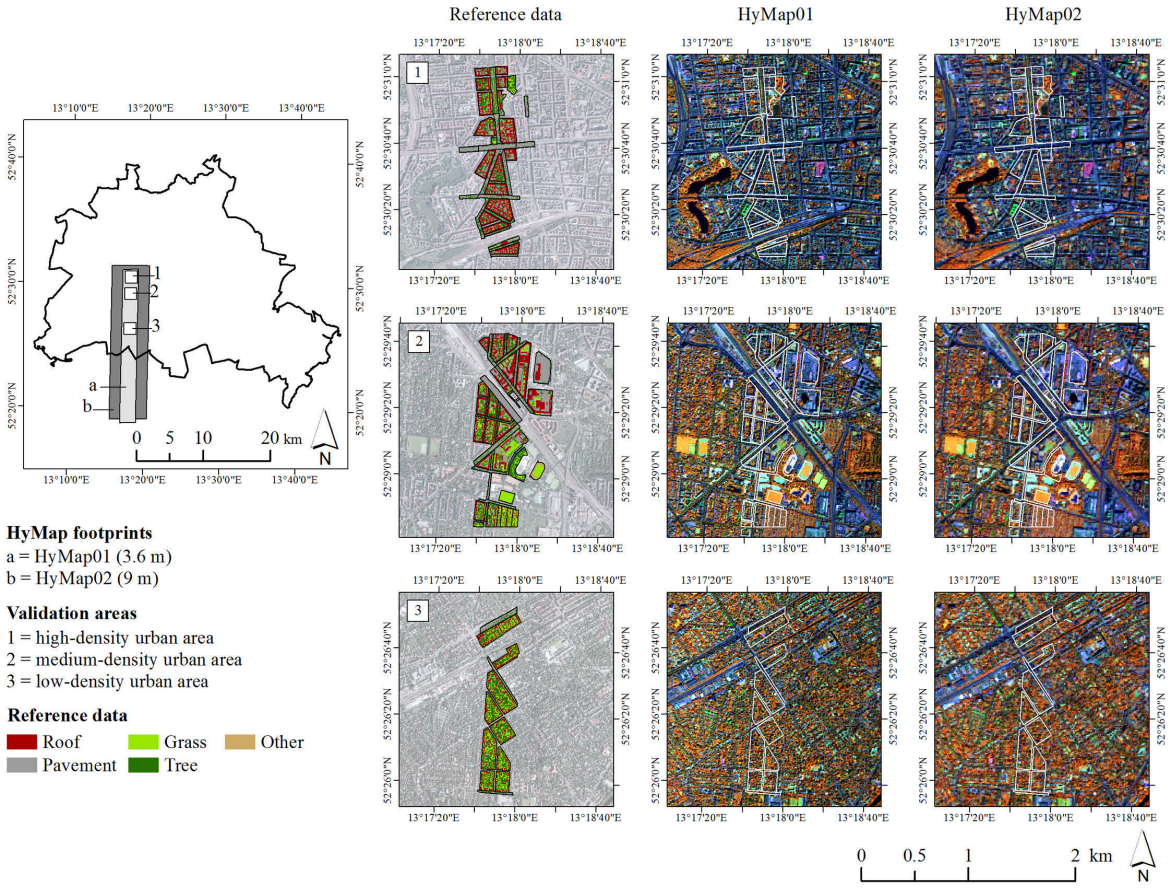


Figure III-1: Study region along Berlin’s urban-rural gradient. The high resolution reference data and the HyMap images (R = 833 nm, G = 1652 nm, B = 632 nm) for the three validation areas are illustrated (polygons indicate the urban blocks used for validation).

Table III-1: Categorized image spectral library of urban materials.

Urban Category	Surface Material	No. Spectra
Roof	Red clay tile	4 (a)
	Red cement tile	3 (a)
	Bitumen	5 (b)
	Brown roof tile	1
	Brown roof shingle	1
	White roof material (polyethylene)	1
	White roof material (unknown)	1
	Zinc roof material	1
Pavement	Asphalt	4 (a, b)
	Concrete	2 (b)
Grass	Grass (intensively manicured)	2 (a)
	Grass (extensively manicured)	1
	Grass (dry)	2 (a)
Tree	Deciduous tree	7 (a)
Other	Tartan (sports ground)	1
	Railtrack (concrete sleepers)	1
	Railtrack (wooden sleepers)	1
	Sand (playground)	1
	Soil	1
	Water	1

Multiple spectra per material considered to account for spectral variability due to variations in (a) illumination and shading effects; and (b) material condition.

(Schläpfer and Richter 2002) (Figure III-1). Pre-processing to reflectance values encompassed system correction (Cocks et al. 1998) and radiometric correction (Richter and Schläpfer 2002). The number of spectral bands was reduced from 128 to 111 in order to limit the influence of noise during the analysis. Due to the nearly similar acquisition time around solar noon, the high sub-pixel accuracy achieved during ortho-rectification and the identical radiometric and atmospheric pre-processing, consistency between both images was given for a comparative study.

2.3 Spectral library

We used a library composed of 41 urban material spectra extracted from HyMap01 as database for generating synthetically mixed training data (Table III-1). Detailed information on the development of the image spectral library was provided in Okujeni et al. (2013), where the library was demonstrated to include relevant spectra necessary for quantifying urban land cover within the same study region. The library includes a representative standard spectrum per material, which was complemented in case of high spectral variability. Complemented spectra account for variations in brightness caused by illumination or shading effects or variations due to differences in material condition.

Spectra were categorized into the four urban land cover categories of interest, *i.e.*, *roof* (17 spectra), *pavement* (6), *grass* (5) and *tree* (7). An additional *other* (6) category, which did not constitute a category of interest in this work, was used to assign surface types with marginal spatial occurrence (e.g., water, sport grounds covered with artificial turf, etc.).

2.4 Reference data

We performed a polygon-wise evaluation of results, a common strategy to statistically assess the accuracy of fraction maps derived from urban remote sensing data (Powell et al. 2007; Roberts et al. 2012). We used urban blocks from the Berlin Urban Environmental Information System (SenStadt 2014b) (Figure III-1), which constitute a suitable spatial unit for many urban environmental applications (Pauleit and Duhme 2000; Heiden et al. 2012). Within the three validation areas, block-wise mean fraction values for a total number of 92 urban blocks were calculated from an available high resolution reference land cover map (Okujeni et al. 2013). To guarantee a representative heterogeneous validation data set, we selected 45 building blocks of different urban structure types, 37 street polygons including roads of different sizes, and 10 green spaces consisting of parks, allotment gardens or sport grounds. All urban block polygons were selected within a ±10° off-nadir region to minimize effects of reflectance anisotropy (Schiefer et al. 2006) and urban 3D-geometry (van der Linden and Hostert 2009) during validation. This way, spectral effects influenced almost entirely the accuracy of fraction maps.

3 Methods

3.1 Synthetically mixed training data

We followed the concept of synthetically mixed training data for regression modeling (Okujeni et al. 2013). The general idea is to use a spectral library to construct a single quantitative training data set for a given land cover category of interest, a so-called target category. Consisting of pairs of spectra and related mixing fractions, this set is subsequently used as input for regression model training. A fraction map of the target category is derived by applying the model to the image data. The subsequent analysis of several target categories allows for a comprehensive mapping of urban land cover.

While generating synthetically mixed training data, we assume a linear mixing process between materials. It cannot be ruled out that nonlinear mixing effects caused by multiple

scattering processes between materials (Roberts et al. 1993; Borel and Gerstl 1994) may affect mapping results. Yet, quantitative effects of nonlinear mixing on the accuracy of fraction estimates are poorly understood (Somers et al. 2011). In the context of urban land cover assessments, linear mixing systematic was frequently adopted and led to valid results (Small 2003; Powell et al. 2007; Franke et al. 2009; Roberts et al. 2012).

We generated synthetically mixed training data as follows: For each of the four categories of interest, the spectral library was partitioned into (i) a target category (e.g., *tree*) and (ii) a background category with all remaining categories (e.g., *roof*, *pavement*, *grass*, *other*). Subsequently, linear mixtures between target and background category spectra were calculated using mixing steps of 20%. This step size was shown to provide sufficient continuous information for quantifying urban land cover using SVR (Okujeni et al. 2013). This resulted in a set of pure spectra, synthetically mixed spectra and related mixing fractions of 0%, 20%, 40%, 60%, 80% and 100% of the respective target category. Spectra within the *other* category were only used as background. Accordingly, synthetically mixed training data were not derived for the *other* category and corresponding surface types, such as water or sports grounds covered with artificial turf, were not explicitly mapped.

To account for the mixing complexity, we constructed two different training data sets for each target category (Figure III-2). Within the first type of training data (SyMix01), we included all possible binary mixtures between target category spectra and background

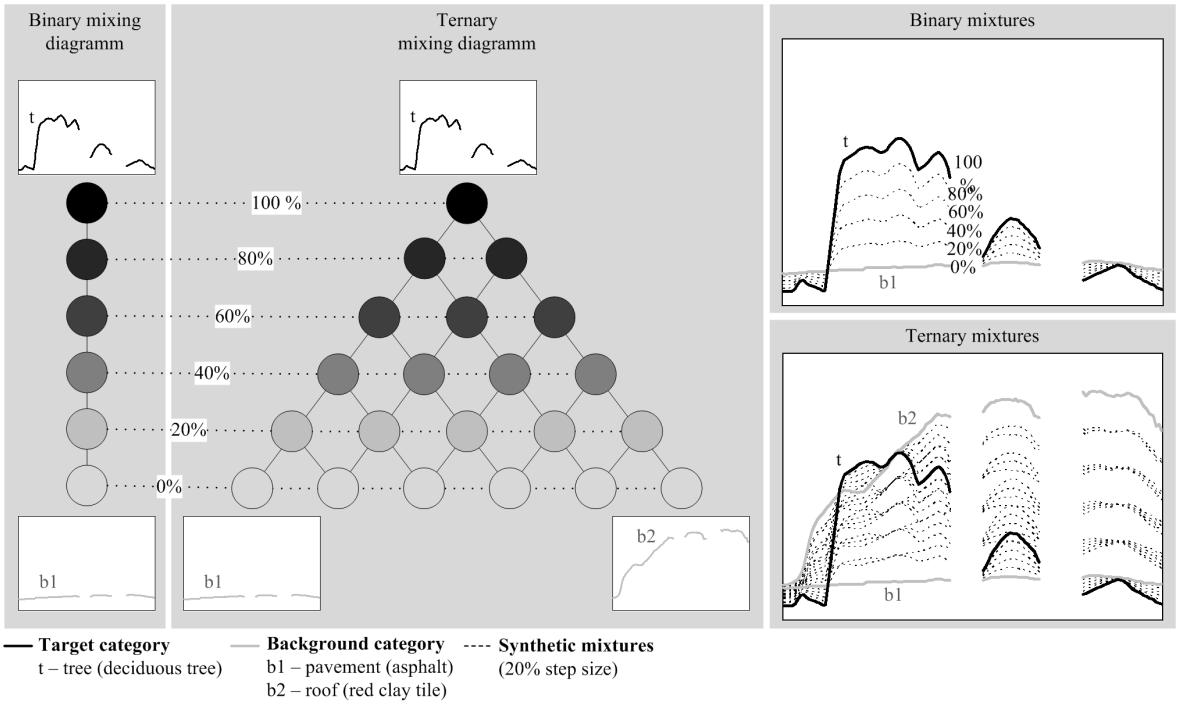


Figure III-2: Generation of binary and ternary synthetic mixtures.

Table III-2: Number of training samples per target category used for regression modeling.

Training Data Set	*Roof*	*Pavement*	*Grass*	*Tree*
SyMix01	1673	881	761	993
SyMix02	3092	1970	1202	1254

category spectra. Material combinations within a category were not included. SyMix01 is considered as training data of lower complexity, as only mixtures of two materials were integrated. Within the second type of training data (SyMix02), we included all possible binary mixtures and selected ternary mixtures between target category spectra and background category spectra. In our case, ternary mixtures were calculated as two-fold binary mixtures and included (i) combinations between a target spectrum (mixing fraction of 100%) and a binary mixed background spectrum (mixtures of two background spectra in 20% increments but treated as having a mixing fraction of 0%) or (ii) combinations between a binary mixed target spectrum (mixtures of two target spectra in 20% increments but treated as having a mixing fraction of 100%) and a background spectrum (mixing fraction of 0%). Knowledge on relevant ternary mixtures was derived from visual inspection of digital orthophotos. We considered both combinations between different categories (e.g., “deciduous tree-red clay tile-asphalt” or “deciduous tree-grass-soil”) and combinations within categories (e.g., “deciduous tree-red clay tile-bitumen” or “asphalt-concrete-tree”). This selective approach was study area specific, however, guaranteed the inclusion of higher order mixtures while at the same time prevented an excessive amount of training samples due to the variety of possible combinations. SyMix02 is considered as training data of higher complexity, as mixtures of two and three materials were integrated. The sample sizes of the eight training data sets, which are each used separately to train one regression model per target category and mixing complexity, are reported in Table III-2.

3.2 Support vector regression (SVR)

Emanating from the field of kernel-based machine learning methods, SVR has been established in remote sensing research, mainly as a powerful, nonlinear technique for quantifying biophysical/chemical plant properties (Camps-Valls et al. 2006; Yu et al. 2013). Details on the underlying concepts of SVR can be found in Schölkopf and Smola (2002). In general, SVR estimates the linear dependency between pairs of n-dimensional input vectors and 1-dimensional target variables by fitting an optimal approximating hyperplane to a set of training samples. The optimal hyperplane is defined by a linear regression model that is found by solving a convex optimization problem. The most common SVR formulation thereby minimizes Vapnik’s ε-insensitive cost function.

Embedded into a kernel framework, SVR is capable of coping with nonlinear data distributions. By using a kernel function, training samples are mapped into a higher dimensional feature space, which is nonlinearly related to the original space and wherein the new data distribution enables a better fitting of a linear function.

The SVR analysis of this work was carried out with imageSVM (Rabe et al. 2010a), which uses the Gaussian kernel function. Parameterization of an SVR required the selection of the kernel parameter σ, the regularization parameter C and the loss function parameter ε. Tuning of the three parameters was carried out via grid search using a cross validation strategy to avoid overfitting.

3.3 Kernel ridge regression (KRR)

KRR has been lately introduced for remote sensing applications (Verrelst et al. 2012; Caicedo et al. 2014). Detailed information on the theoretical background of KRR can be found in Hastie et al. (2009). Similar to SVR, KRR maps the training samples into a higher dimensional feature space by making use of a kernel function. There, KRR minimizes the squared residuals, and can be therefore considered as the kernel version of the regularized least squares regression (Verrelst et al. 2012).

The KRR analysis of this work was carried out with the ARTMO software package (Caicedo et al. 2014). KRR modeling required the tuning of two parameters, i.e., the Gaussian kernel function parameter σ and the regularization parameter C. Parameter optimization was carried out via standard cross validation.

3.4 Neural network regression (NN)

Artificial neural networks were one of the first non-parametric, nonlinear techniques to be used for remote sensing applications (Bacour et al. 2006; Van de Voorde et al. 2009; Cernicharo et al. 2013). Details on the theoretical concepts of NN can be found in Haykin (1999). A NN is a model that establishes the relationship between input vectors and output variables through a connected structure of neurons organized in layers. Each neuron basically performs a linear regression followed by a nonlinear function. Neurons of different layers are interconnected with the corresponding links, i.e., weights. NN modeling requires the selection of a NN structure, the initialization of weights, shape of nonlinearity and learning rate as well as the regularization of parameters to avoid overfitting (Verrelst et al. 2012).

The NN analysis of this work was carried out with ARTMO (Caicedo et al. 2014), which uses a fully connected standard multi-layer perceptron and a hyperbolic tangent as nonlinear activation function. We selected just one hidden layer of neurons, weights were randomly initialized and the NN structure was optimized using a squared loss function. Model parameterization was carried out via a standard cross validation procedure.

3.5 Random forest regression (RFR)

Random forest (Breiman 2001), an ensemble method based on multiple decision trees, has been used in a variety of remote sensing studies (Chan and Paelinckx 2008; Walton 2008). Decision trees are non-parametric predictive models, which are based on a multistage decision scheme. They exist in the form of classification and regression trees, as they are suited to predict both discrete and continuous output variables. The tree structure consists of a root node, internal nodes (splits) and terminal nodes (leaves). At each split and until a leaf node is reached, a decision rule is applied to partition the training data recursively into increasing number of smaller homogenous subsets. A target variable is assigned to each sample according to the leaf node (Breiman et al. 1984; Friedl and Brodley 1997). As an ensemble method, RF combines the results of many individual decision trees in order to improve the overall prediction performance. Independent decision trees are created through bagging (bootstrap aggregation), where each decision tree uses a random sample (with replacement), and through random feature selection at each split node. Training samples that are excluded, so-called “out-of-bag” (OOB) samples, can be used for independent validation. For the regression case, the final prediction is obtained by averaging the results of individual trees (Breiman 2001).

The random forest regression analysis was carried out using imageRF (Waske et al. 2012). The number of trees was set to 100, the ratio of bootstrap data to OOB data was 2/3–1/3 and the number of randomly selected features corresponded to the square root of all features. For each tree, the respective OOB data was used for predicting and the accuracy of combined results was assessed using the OBB error. The learning curve of the OOB error with increasing number of trees was used to verify whether reliable parameters were set.

3.6 Partial least squares regression (PLSR)

Partial least squares regression (Wold et al. 2001) has been widely used for quantifying vegetation properties by means of hyperspectral data (Smith et al. 2003; Schmidtlein et al.

2012; Yu et al. 2013). PLSR is a multivariate method that is very similar to principle component regression, where the dimensionality of the input features is first reduced with respect to their information content followed by multiple linear regression modeling. Unlike principle component regression, where only correlations within the input feature space are considered, PLSR transforms the input features into a few uncorrelated latent vectors, which are generated not only with respect to their variance but also with respect to their explanatory power to predict a response variable during linear regression. Latent vectors are statistically independent linear combinations of the original input features. PLSR is well suited for handling highly correlated data (Wold et al. 2001; Feilhauer et al. 2010).

The PLSR analysis of this work was carried out using autopls (Schmidtlein et al. 2012). autopls includes an iterative backward selection of latent vectors, which reduces the number of predictors in the regression following Westad and Martens (2000) and Chong and Jun (2005). A cross-validation strategy was used to prevent overfitting during model training. The final model was selected based on the learning curve during backward selection, i.e., the model with the smallest training error was used for regression modeling on the image.

3.7 Validation

Validation of results was carried out by comparing modeled versus reference fractions at urban block scale (block-wise mean values). We calculated both class-wise accuracy scores (performance measures for a single fraction map of a target category) and average accuracy scores (mean of class-wise accuracies over all categories). We used the mean absolute error (MAE) and the root mean square error (RMSE) as measures of accuracy. MAE and RMSE are defined as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad (2)$$

where y_i are the modeled fractions, x_i are the reference fractions and n the number of validation samples. To evaluate the goodness of fit between estimated and reference fractions, we used the coefficient of determination (R^2), given by

$$R^2 = \frac{(\sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y}))^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

where $\bar{y}$ and $\bar{x}$ are the averages of the modeled and reference fractions, as well as the slope and the intercept of a fitted least squares linear regression model.

4 Experimental setup and results

4.1 Experimental setup

To answer the research questions, we carried out three experiments using the two HyMap images (HyMap01, HyMap02), synthetically mixed training data at different complexities (SyMix01, SyMix02) and five different regression techniques (SVR, KRR, NN, RFR and PLSR). The experiments are summarized as follows:

- Exp. 1: We investigated the efficiency of the regression techniques to utilize synthetically mixed training data for quantifying urban land cover. For each regressor and each target category, we trained one regression model using SyMix01. Subsequently, we applied the models to HyMap01 to derive fraction maps at 3.6 m spatial resolution.
- Exp. 2: We investigated the influence of coarser resolution image data on the quality of model predictions. We therefore applied regression models used in Exp. 1 to HyMap02 to derive fraction maps at 9 m spatial resolution.
- Exp. 3: We investigated whether more complex synthetic mixtures help to improve fraction estimates from the coarser resolution data. For each regressor and each target category, we trained one regression model using SyMix02. Subsequently, we derived a second set of fraction maps at 9 m spatial resolution by applying models to HyMap02.

4.2 Average and class-wise accuracies of experimental results

Average accuracy scores of urban land cover fraction estimates are reported in Table III-3. In Exp. 1, both kernel methods SVR and KRR show best performances, with high accuracies (e.g., MAE below 12%, R^2 above 0.75) at 3.6 m spatial resolution. Results of NN, RFR and PLSR are less accurate. In Exp. 2, deteriorating mapping accuracies are observed for most of the regressors when applying models to the coarser resolution image. Still, both kernel methods SVR and KRR produce considerably high average accuracies

(e.g., MAE below 13%, R^2 above 0.65) while accuracies of NN, RFR and PLSR remain less accurate. NN produces inconsistent results with opposing average accuracy scores, i.e., improving MAE but also decreasing R^2. In Exp. 3, the increased complexity of the training data particularly improves SVR-, KRR- and PLSR-based estimates. For NN and RFR, only slight improvements are observed. Also, in this experiment, SVR and KRR remain the best performing techniques with high average accuracies (e.g., MAE of 11.5% and below, R2 of 0.71).

Class-wise MAE values (Figure III-3) reveal the performances of the five regressors when mapping different urban surface types. For a more general evaluation, accuracy scores are

Table III-3: Average accuracies of urban land cover fraction maps derived from different regression algorithms for the three experiments. Highest accuracy scores for each experiment are highlighted by bold numbers.

Experiment	Data	Method	MAE (%)	RMSE (%)	R2	Slope	Intercept
Exp. 1	HyMap01 SyMix01	SVR	**9.8**	**12.5**	0.76	**0.62**	**12.5**
		KRR	11.5	13.8	**0.77**	0.60	13.8
		NN	17.9	20.8	0.61	0.49	20.8
		RFR	16.0	18.4	0.53	0.35	18.4
		PLSR	15.8	18.8	0.55	0.46	18.8
Exp. 2	HyMap02 SyMix01	SVR	**12.3**	**15.3**	0.66	0.56	**15.3**
		KRR	12.7	15.5	**0.70**	**0.57**	15.5
		NN	16.8	19.2	0.54	0.46	19.2
		RFR	17.8	20.4	0.50	0.38	20.4
		PLSR	15.9	18.5	0.51	0.41	18.5
Exp. 3	HyMap02 SyMix02	SVR	**10.7**	**13.4**	**0.71**	**0.63**	**13.4**
		KRR	11.5	14.2	**0.71**	0.59	14.2
		NN	16.2	18.5	0.56	0.47	18.5
		RFR	17.1	19.5	0.51	0.39	19.5
		PLSR	13.7	16.7	0.55	0.45	16.7

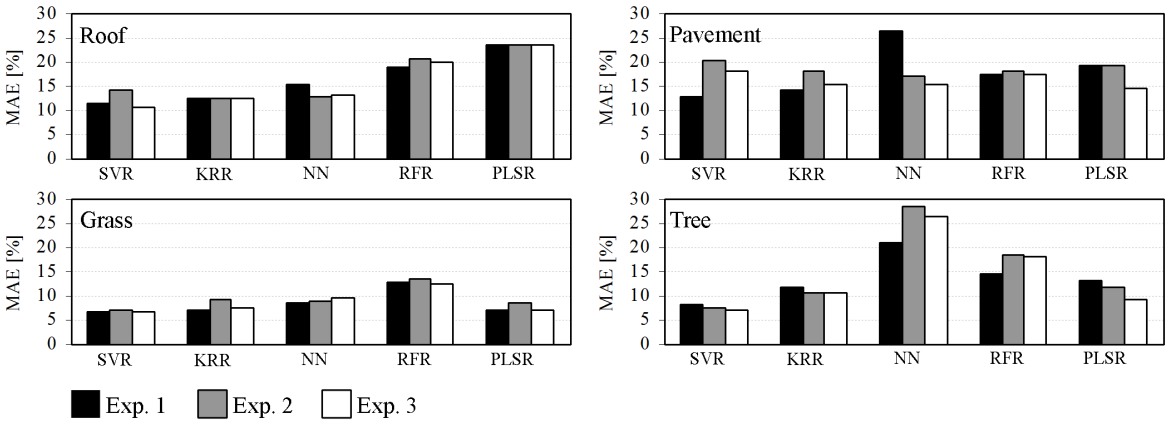


Figure III-3: Class-wise mean absolute error (MAE) of urban land cover fraction maps derived from different regression algorithms for the three experiments.

classified as highly accurate (MAE below 10%), accurate (MAE between 10% and 15%), sub-optimal (MAE between 15% and 20%) and inaccurate (MAE above 20%). In Exp. 1, both kernel methods SVR and KRR produce accurate *roof* and *pavement*, and accurate to highly accurate *grass* and *tree* mappings. The performance of NN is highly variable with sub-optimal roof, highly accurate grass and inaccurate *pavement* and *tree* estimates. The ability of RFR and PLSR to predict roof and pavement is rather low, with sub-optimal to inaccurate mapping accuracies. In contrast, *grass* and *tree* estimates by RFR and PLSR are accurate to highly accurate at 3.6 m spatial resolution.

The impact of the coarser resolution image (Exp. 2) and the training data of higher complexity (Exp. 3) on the quality of fraction maps strongly varies between the regressors and the urban land cover types. Decreasing performances, which are fully compensated for by the more complex training data, are observed for SVR-based *roof* and KRR-based *grass* estimates. Both mappings remain accurate to highly accurate at 9 m spatial resolution. *Pavement* predictions by the kernel methods follow similar trends, however, with only partly compensated accuracies by the more complex training data. This results in degraded, sub-optimal *pavement* estimates by SVR and KRR on the coarser resolved image. The remaining mappings by SVR and KRR are relatively stable with similar accuracies when compared to results at 3.6 m. NN reveals erratic patterns in accuracy scores between the urban categories, *i.e.*, *roof* and especially *pavement* inconsistently improve to accurate and sub-optimal mappings mainly due to the decreased spatial resolution, *grass* predictions remain stable and highly accurate and the inaccurate *tree* mapping further deteriorates. The performance of RFR and PLSR for mapping *roof* at 9 m spatial resolution remains low. *Pavement* estimates by RFR remain sub-optimal, whereas PLSR predictions improve through integrating training data of higher complexity. The performance of RFR and PLSR to map *grass* on the coarser resolution image is still high, with relatively stable accuracies between the experiments. Considering the *tree* category, RFR-based predictions clearly deteriorate at 9 m resulting in sub-optimal performances despite training data of higher complexity. PLSR reveals accurate *tree* predictions at 9 m, with improved accuracy scores compared to the result on the higher resolved image.

The scatterplots (Figure III-4) reveal the regions of agreement and disagreement between reference fractions and modeled urban land cover fractions derived from HyMap02 at 9 m spatial resolution using SyMix02 (Exp. 3). A general representation of how building blocks from different urban densities, street corridors and green spaces influence the accuracies of

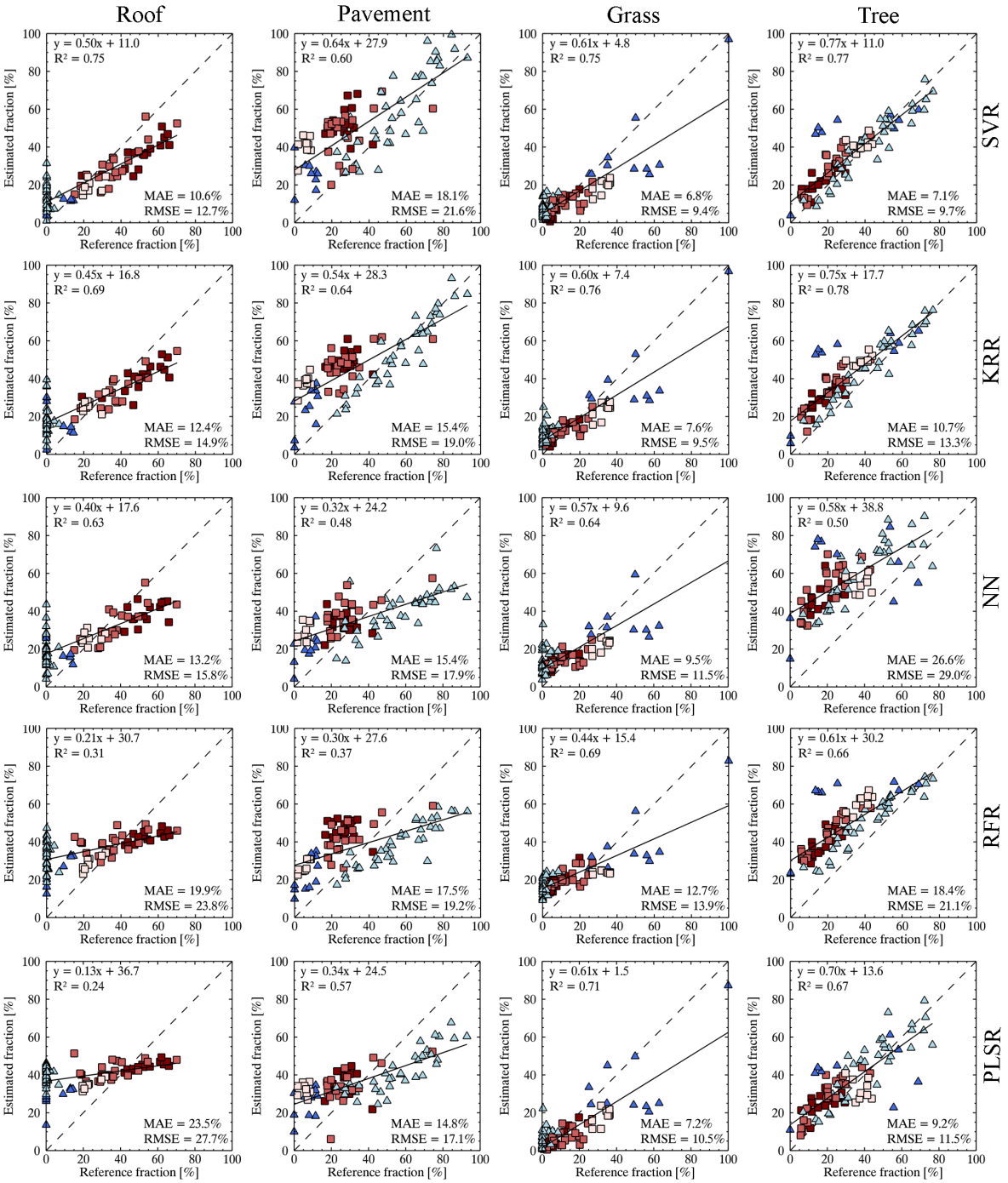


Figure III-4: Scatterplots of *roof*, *pavement*, *grass* and *tree* estimates compared to reference data at urban block scale. Results of Exp. 3 (HyMap02, SyMix02) are presented. The following symbols are used: Squares = building blocks (dark red = high-density-, red = medium-density-, light red = low-density urban area); triangles = green spaces (blue), street polygons (light blue).

fraction estimates is given. The accurate *roof* estimates by SVR, KRR and NN are underlined by the linear data distribution close to the 1:1 diagonal line. Within building blocks, *roof* fractions are precisely mapped in the low- and medium-density urban area and slightly underestimated within the high-density urban area. One common major source of error for all three regression techniques are misleadingly mapped *roof* fractions along streets or within green spaces. The lower performance to map *pavement* by all regressors is explained by a general higher rate of scattering and a characteristic overestimation of fractions within the building blocks of different densities and green spaces. In contrast, *pavement* fractions mapped along street corridors appear accurate, particularly for KRR. *Grass* results are accurate for all regressors. Only RFR shows a slight deviation from this trend. The high accuracies of *tree* estimates by SVR, KRR are underlined by the strong linear correlation with the reference data. PLSR shows similar high performances, however, with increased scattering. The four green spaces, which clearly deviate from the 1:1 diagonal line for both *grass* and *tree* estimates, correspond to allotment gardens. These inaccuracies repeatedly occur for all regressors in all experiments and relate to uncertainties in the reference data, as the manual mapping of highly fragmented vegetation patches within allotment gardens appears critical.

4.3 Quantifying urban land cover at multiple spatial scales using SVR

Summarizing from the tested regression techniques, SVR and KRR yielded the most accurate land cover estimates at multiple scales. We selected SVR for a pixel-wise representation of urban land cover at 3.6 and 9 m spatial resolution (Figure III-5). Results depict the typical structure of Berlin along the urban–rural gradient. The high-density urban area is characterized by a high amount of buildings and paved areas as indicated by the *roof* and *pavement* fraction maps. Apart from a few industrial areas and highways, these densities decrease towards the urban fringe. Yet, reported overestimates in *pavement* fractions within the medium- and low-density urban area are obvious particularly where vegetation, and thus shadowing within vegetation stands, dominate the scenery. High vegetation proportions even within the city center relate to the high amount of green spaces and street trees that characterize the cityscape of Berlin. Nevertheless, a gradual increase in *grass* and *tree* fractions is observed when moving towards the urban fringe. *Other* surface types, such as water bodies in the high-density urban area or sports grounds covered with artificial turf in the medium-density urban area, were only used as background category during processing and appear bright (fractions of 0%) in the maps. Overall, consistent and

similar spatial patterns in fraction values are observed between the urban land cover maps at 3.6 and 9 m spatial resolution.

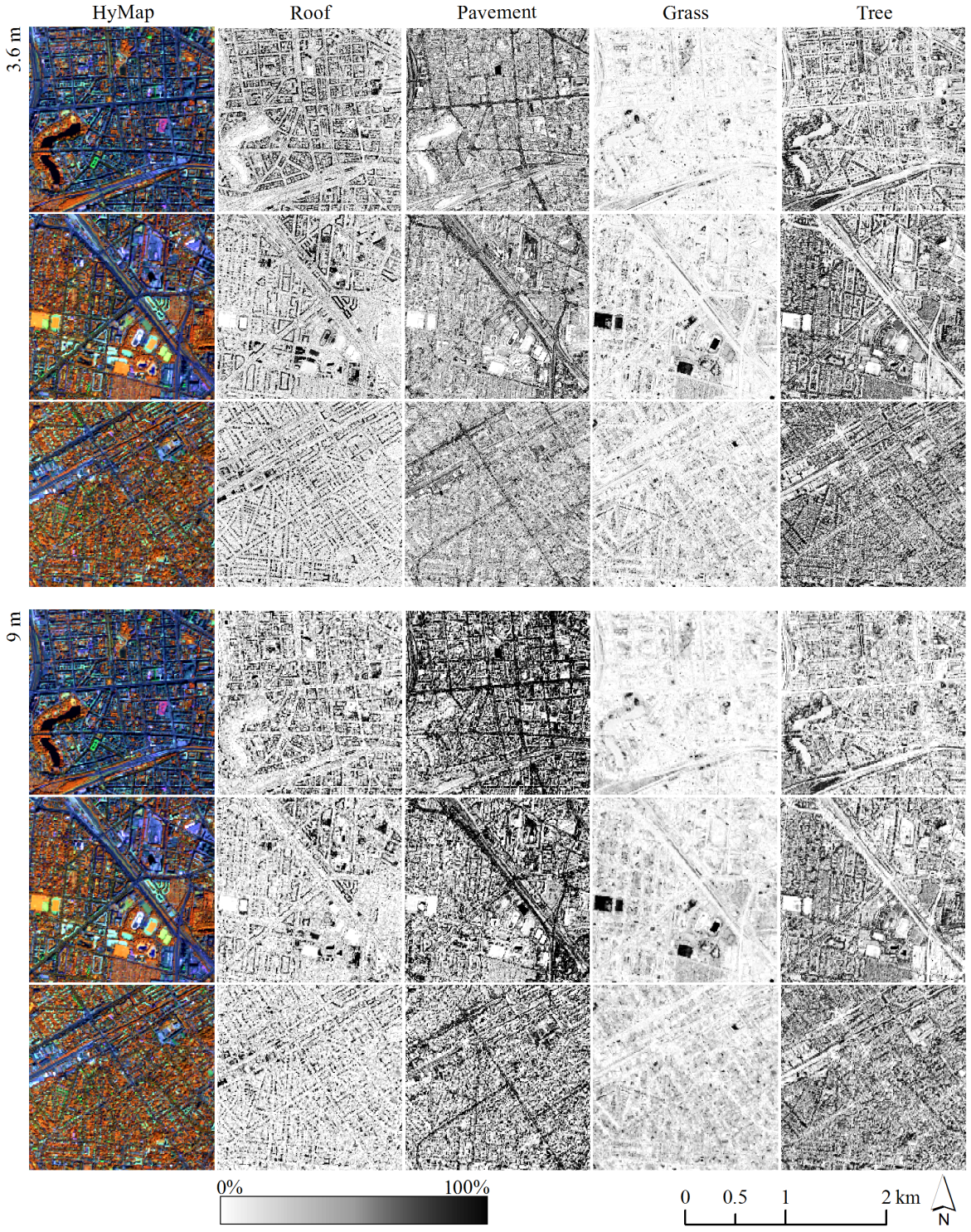


Figure III-5: SVR-based fraction maps of *roof*, *pavement*, *grass* and *tree* at 3.6 m (Exp. 1) and 9 m (Exp. 3) spatial resolution. Fraction maps were linearly stretched between 0% and 100%. For information on the locations of the individual subsets, the reader is referred to Figure III-1.

5 Discussion

We tested the combination of five advanced regression techniques with synthetically mixed training data of different complexities for quantifying urban land cover from HyMap data at 3.6 and 9 m spatial resolution. The use of synthetic spectral mixtures poses a straightforward strategy enabling sub-pixel mapping assessments by means of regression analysis. The approach requires a library with relevant material spectra for the study area. Linking the image spectral library to both images was unproblematic due to the consistency between the data sets. Universal spectral databases may be used alternatively, provided that library spectra and imagery are accurately cross-calibrated. The identical inputs from a spectral library used during model training guaranteed a fair methodological comparison. Block-wise accuracy assessment with an independent reference map gave an important insight into the generalization capabilities of the regressors when applied to a real data scenario.

We considered the four thematically meaningful and spectrally challenging urban land cover types *roof*, *pavement*, *grass* and *tree*. The *roof* category is a typical example for a spectrally complex class, which is characterized by a multi-modal spectral distribution due to the high number of different materials at different physical conditions and compositions (Herold et al. 2004; Heiden et al. 2007). The *pavement* category is a typical example for a class that is highly subject to spectral similarities. This is due to a low number of spectrally distinct materials, most of them with very similar spectral signatures when compared to dark spectra from background cover types (e.g., street asphalt *vs.* bituminous roofing materials) (Herold et al. 2004; Heiden et al. 2007). Likewise, high spectral similarity between vegetation types impedes the differentiation of *grass* and *tree* in such a heterogeneous setup. The spectral complexity of urban materials is further amplified by illumination and shadowing effects. Particularly shaded areas have been reported to influence the accuracy of urban land cover maps derived from hyperspectral imagery (Roessner et al. 2001; van der Linden and Hostert 2009; Okujeni et al. 2013). By choosing the four urban categories, the ability of the tested regression methods to deal with spectral complexity, similarity and variability was explored.

Two major findings can be drawn from Exp. 1, where HyMap data at 3.6 m spatial resolution was used to quantify urban land cover. First, we could once more demonstrate that the concept of synthetically mixed training data enables the use of empirical regression

techniques for sub-pixel mapping of single urban land cover categories. Second, we could demonstrate that the quality of fraction maps strongly depends on the selected regression method. Both kernel-based regression techniques SVR and KRR proved to be well suited for quantitatively analyzing impervious and natural urban sub-categories using hyperspectral data. High average accuracies (Table III-3) point to the ability of SVR and KRR for a comprehensive quantification of urban land cover. Class-wise accuracies (Figure III-3) indicate the efficiency of SVR and KRR to discriminate the challenging urban categories *roof*, *pavement*, *grass* and *tree* with a low degree of confusion caused by spectral complexity and similarity.

In Exp. 1, SVR and KRR outperformed NN, RFR and PLSR in most cases. They revealed either erratic (NN) or limited (RFR and PLSR) performances, which makes a general evaluation on their use for comprehensively mapping urban land cover from the 3.6 m HyMap data difficult. Yet, all three methods illustrate potential for mapping individual urban categories, e.g., *roof* and *grass* by NN and *grass* and *tree* by KRR and PLSR, respectively.

The change in spatial resolution from 3.6 m to 9 m implies a loss of spatial detail and, even more important, a significant increase in the number of mixed pixels and in the abundance of different materials contributing to the mixed signal. Overall, both kernel methods SVR and KRR yielded most accurate results on the coarser resolution image. Regression models proved to be stable with regard to the spatial and spectral differences between the two images (Exp. 2) and effectively utilized the additional information provided by more complex ternary mixtures to compensate losses in mapping quality (Exp. 3). The integration of ternary mixtures in addition to the binary mixtures always appears worthwhile, as no negative impacts were observed. These findings are supported by the average accuracy scores (Table III-3), which only slightly degrade between Exp. 2 and Exp. 1 (ΔMAE = −2.5% for SVR, ΔMAE = −1.2% for KRR) and clearly improve when comparing Exp. 3 to Exp. 1 (ΔMAE = −0.9% for SVR and 0.0% for KRR). *Roof*, *grass* and *tree* can be mapped at similar high precision by SVR and KRR at 9 m when compared to results at 3.6 m spatial resolution (Figure III-3). However, the precise differentiation of *pavement* from coarser resolution data is challenging. Degraded sub-optimal *pavement* accuracies obtained by both kernel methods likely relate to the dominance of spectrally similar materials, which become increasingly indistinct and difficult to differentiate at 9 m.

Also in Exp. 2 and 3, SVR and KRR mostly showed superior mapping performances over NN, RFR and PLSR. The suitability of NN to reliably cope with the coarser resolution HyMap data when mapping individual urban land cover categories is ambiguous. For example, *roof* and particularly *pavement* estimates are improved to make comparable accuracies to those obtained by both kernel methods. However, this finding is controversial as the enhancement is rather related to the increased pixel size than to the integration of more complex training signatures. Lower performances by RFR and PLSR to map *roof* and *pavement* on the coarser resolution data question the use of both methods when mapping the spectrally complex or similar impervious sub-categories, while both methods appear suitable for mapping urban vegetation types.

Results of this comparative study demonstrate that SVR and KRR yielded the most accurate and reliable estimates from HyMap data at 3.6 and 9 m spatial resolution, whereas NN, RFR and PLSR revealed some major limitations. The quality of fraction maps (Figure III-5) to depict the typical structure along Berlin's urban-rural gradient underscores the suitability of SVR to be used along with synthetically mixed training data for quantifying urban land cover at multiple spatial scales. Nevertheless, several aspects appear worthwhile for discussion. First, our regression techniques were regularly adopted in the remote sensing context (Schmidtlein et al. 2012; Verrelst et al. 2012; Okujeni et al. 2013; Caicedo et al. 2014), however, caution must be exercised that different implementations of these algorithms exist. Second, the strategy to consider only selected ternary mixtures using expert knowledge prevented an excessive amount of training samples. Higher order mixtures within the image data that were not accounted for during training may affect the quality of mappings. Yet, achieved accuracies are valid and the efficiency of the regressors to utilize more complex training mixtures was well explored. Third, results demonstrate that even both the best performing kernel methods did not completely overcome well known problems arising from spectral similarities or shadowing. Particularly, the unambiguous quantification of rooftops or paved-areas was not completely possible and uncertainties in fraction maps remain. Typical examples are misleadingly mapped *roof* fractions along streets or within green spaces and overestimated *pavement* fractions within most building blocks and green spaces (Figure III-4), caused by the spectral similarity between dark roofing materials, street asphalt and canopy shade. The use of multi-sensor approaches that take advantage of information beyond the spectral domain may help to overcome such deficiencies (Heiden et al. 2012). Here, we fully focused on the opportunities and limitations of hyperspectral information for assessing urban land cover.

6 Conclusions

Exploring the applicability of powerful quantitative methods is crucial for the development of robust, universal modeling approaches that are needed once imaging spectrometer data from spaceborne hyperspectral sensors will be available. In this paper, we investigated the potential of five advanced regression approaches combined with synthetically mixed training data for quantifying urban land cover using HyMap data at 3.6 and 9 m acquired over Berlin, Germany. The universality of the approach for urban sub-pixel mapping was tested by exploring (1) the efficiency of the different regressors to utilize synthetically mixed training data for producing accurate urban land cover maps; (2) the impact of coarser resolution data on the quality of model predictions; and (3) the need to adapt the training data to the mixing complexity of the image. In most cases, both kernel-based approaches SVR and KRR showed superior average and class-wise mapping performances when compared to NN, RFR and PLSR. Average MAE values show that (1) SVR and KRR outperform the remaining regressors by at least 4.3% at 3.6 m spatial resolution; that (2) for SVR and KRR average MAE values increase by 2.5% and 1.2%, respectively, when decreasing the spatial resolution from 3.6 m to 9 m; and that (3) the use of more complex models improves average MAE values by 1.6% and 1.2% for SVR and KRR.

Based on the findings of this case study, we conclude that the use of synthetically mixed training data is well suited for quantifying spectrally challenging urban land cover types by means of empirical regression modeling. The problems in setting up such approaches as mentioned in the introduction are therefore overcome. At the same time, the general strength of kernel regression methods for urban land cover assessments is illustrated, which had been extensively demonstrated for per-pixel classification problems. Such regression approaches can, hence, be incorporated into quantitative mapping. SVR and KRR models proved to be stable with regard to the amplified mixing scenario and well utilized the added value of more complex training mixtures to cope with deficiencies of coarser resolution data. However, urban land cover assessments on a purely spectral basis will always bear limitations and specific challenges such as spectral similarities between land cover types or confusion caused by shadowing will always remain.

The combination of advanced regression techniques with synthetically mixed training data is straightforward, repeatable and relies on the availability of a spectral library relevant for the study area. Compared to per-pixel classifiers, the proposed workflow relativizes the mixed pixel problem typical for urban remote sensing data. This is important for future

studies, where the reliability, validity and general transferability of the approach on image data from different sensor types and different urban areas will be tested.

Acknowledgments

This research was funded by the German Research Foundation (Project No. HO 2568/2-2) and partly funded as part of the EnMAP Core Science Team activities (FKZ 50EE0949) by the Federal Ministry of Economics and Technology (BMWi). We thank the five anonymous reviewers for their valuable comments. The authors are grateful to the German Aerospace Centre (DLR) for data acquisition and pre-processing during the HyEurope 2009 Campaign and the Berlin Senate Department for Urban Development for providing data from the Urban and Environmental Information System.

Chapter IV:
Extending the Vegetation-Impervious-Soil model using simulated EnMAP data and machine learning

Remote Sensing of Environment (under review)[*]

Akpona Okujeni, Sebastian van der Linden and Patrick Hostert

[*] A revised version of this chapter was meanwhile published in: Okujeni, A., van der Linden, S., & Hostert, P. (2015). Extending the vegetation-impervious-soil model using simulated EnMAP data and machine learning. *Remote Sensing of Environment*, 158, 69-80.

Abstract

The upcoming hyperspectral satellite mission Environmental Mapping and Analysis Program (EnMAP) will provide timely and frequent imaging spectrometer data globally. This will create unprecedented opportunities for a variety of environmental research fields and lead to manifold novel applications. These opportunities specifically apply to challenging environments, including heterogeneous urban landscapes. In this paper, we explored the potential of EnMAP data for quantifying land cover along the urban-rural gradient of Berlin, Germany. Land cover fraction maps from a simulated EnMAP scene at 30 m spatial resolution were derived based on support vector regression (SVR) combined with synthetically mixed training data. Results demonstrate that EnMAP imagery will be well suited for mapping *impervious*, *vegetation* and *soil* surface types according to the VIS framework. Moreover, EnMAP data will allow extending the VIS framework by more detailed sub-categories such as *roof* and *pavement*, or *low vegetation* and *tree*. We advice caution, though, that also spaceborne imaging spectrometer data of improved quality will not completely overcome well known phenomena of spectral similarity between materials or spectral confusion caused by the presence of shaded areas. To identify possible benefits and limitations of EnMAP data, comparisons to fraction maps derived from a higher resolution hyperspectral HyMap image at 9 m spatial resolution and a multispectral Landsat-like image at 30 m spatial resolution were drawn. First, we demonstrate that both VIS and extended VIS mapping reveal similar accuracies compared to maps from spatially higher resolution data. Second, we illustrate the superiority of the higher spectral information content for improved and extended urban land cover mapping compared to multispectral data. Overall, this study provides important insights into the potential of spaceborne imaging spectrometer and specifically future EnMAP data for urban remote sensing.

1 Introduction

Urban areas are dynamically changing environments and subject to population growth or decline, followed by new formations or reconfigurations of residential areas, commercial and industrial zones, traffic infrastructures and green spaces (Antrop 2004; Besussi et al. 2010). Changes in urban land cover and land use along with the concentration of human consumption, production and waste discharge have led to local to global ecological concerns (Grimm et al. 2008). Enhancing our understanding of how urbanization impacts the environment is therefore of great importance, particularly with regard to the development of more sustainable urban pathways (Alberti 2005; Grimm et al. 2008; Pickett et al. 2011).

The urban-rural gradient approach has proven to be a suitable concept to study interactions between urban development and ecological patterns and processes (McDonnell and Pickett 1990; Blair 1996), and thus to develop a more thorough understanding of the coupled human-natural urban system. The urbanization component along the gradient can be quantified by different measures, including population statistics, physical surface properties or landscape metrics (Hahs and McDonnell 2006). By providing synoptic views in space and time, optical remote sensing data is ideally suited for mapping and monitoring the composition and the spatial patterns of land cover along urban-rural gradients. Land cover data or spatial metrics derived from satellite imagery therefore often serve as measures of urbanization within urban environmental research (Hahs and McDonnell 2006; Kong and Nakagoshi 2006; Yu and Ng 2007).

Several broad- and fine-scale land cover categorization frameworks have been developed for urban remote sensing: The VIS model (Ridd 1995) is a universal, globally applicable mapping scheme that quantifies urban land cover into the three components “vegetation”, “impervious” and “soil”. Especially the mapping of impervious surfaces (a comprehensive review was provided by Weng (2012)) and urban vegetation abundance (Weng et al. 2004; Small and Lu 2006) has received much attention in the remote sensing community. Corresponding maps serve to measure the degree of urbanization (Hahs and McDonnell 2006; Zhou and Wang 2011) or can be directly related to urban environmental conditions such as the urban heat island (Weng et al. 2004; Imhoff et al. 2010). The SVD model (Small and Lu 2006) constitutes an alternative framework quantifying urban areas into the biophysically determined reflectance properties high albedo “substrate”, “vegetation” and

"dark surface". The SVD scheme is a robust, universal modeling approach that is particularly useful to derive meaningful estimates of urban vegetation (Small 2001, 2003; Small and Lu 2006). Besides the VIS and SVD frameworks, different hierarchically structured categorization schemes have been developed to assess the fine-scale heterogeneity of urban areas. These schemes further differentiate impervious, pervious or vegetation sub-categories up to individual urban surface materials and vegetation species (Roessner et al. 2001; Herold et al. 2003; Heiden et al. 2007; Franke et al. 2009). The differentiation of such detailed urban surface properties plays an important role in enhancing our understanding of ecological processes that are for example related to urban climate (Gluch et al. 2006; Grimmond 2007; Middel et al. 2014), hydrology (Arnold and Gibbons 1996; Pauleit and Duhme 2000; Sjöman and Gill 2014) or biodiversity (Blair 1996; McKinney 2002; Fontana et al. 2011).

Globally available multispectral satellite data, especially from the Landsat platforms, have been extensively exploited for mapping the VIS components of different cities (Rashed et al. 2003; Wu and Murray 2003; Powell et al. 2007). However, many urban materials are not spectrally distinct when represented by broadband reflectance data. This often leads, for example, to difficulties in accurately distinguishing different impervious surface types and pervious soils (Herold et al. 2003; Small and Lu 2006; Powell et al. 2007). These problems can be circumvented by mapping the SVD surface reflectance components. Nevertheless, linking spectral information with more thematically-oriented land cover information, i.e., beyond the brightness level of non-vegetated surface types, remains desirable. In contrast, hyperspectral data – also referred to as imaging spectrometer data – enhances the separability of surface materials by providing continuous spectral information that resolves material-specific absorption features (Herold et al. 2004; Heiden et al. 2007). This generally enables the use of fine-scale categorization frameworks for urban land cover assessments, including material- or species-oriented mapping (Herold et al. 2003; Franke et al. 2009; Alonzo et al. 2014) or mapping of thematically detailed surface types such as rooftops, paved-areas, pervious surfaces, grass- and tree-cover (Roessner et al. 2001; van der Linden and Hostert 2009).

Compared to widely available multispectral satellite imagery, high quality imaging spectrometer data sets are so far almost exclusively constrained to spatially limited, high spatial resolution scenes acquired from airborne sensors. Only few studies exploited urban images acquired by the spaceborne Hyperion instrument (Cavalli et al. 2008; Weng et al. 2008). Amongst others, this may be explained by the reduced instrumental performance of

Hyperion compared to airborne hyperspectral sensors (Kruse et al. 2003; Roberts et al. 2003). However, the drawback in data availability is likely to change in the near future. Several hyperspectral spaceborne sensors are currently planned, including the Environmental Mapping and Analysis Program (EnMAP; Kaufmann et al. 2008), the Hyperspectral Infrared Imager (HyspIRI; Roberts et al. 2012), the Hyperspectral Precursor of the Application Mission (PRISMA; Galeazzi et al. 2008) or the Hyperspectral Imager Suite (HISUI; Matsunaga et al. 2013). These systems will provide imaging spectrometer data globally, which will create unprecedented opportunities for novel environmental applications. Given the advantages of hyperspectral information for describing detailed surface properties, this may contribute to the increased availability of improved, thematically detailed and regular updated urban land cover maps at local, regional or global scales.

Despite these promising technological developments, deriving accurate maps of urban surface types from spaceborne imaging spectrometer data will pose a great challenge. On the one hand, urban land cover assessments must cope with the vast spectral diversity of urban materials and related characteristics such as high within-class variability or between-class similarity (Herold et al. 2004; Heiden et al. 2007; van der Linden et al. 2007). On the other hand, image acquisitions from space will bring along coarser spatial resolution compared to the currently most common airborne data collections. This results in a loss in spatial detail and a significant increase in mixed pixels, specifically regarding spatially heterogeneous urban environments. Powerful quantitative approaches that are capable of coping with spectral complexities and spectral mixtures are therefore needed to fully exploit the latent information content of imagery from future spaceborne imaging spectrometers. In this context, several sub-pixel mapping methods have been suggested, including spectral unmixing or regression. Multiple endmember spectral mixture analysis (MESMA; Roberts et al. 1998) is a widely used spectral unmixing technique that has been successfully adopted for quantifying urban land cover (Powell et al. 2007; Franke et al. 2009; Roberts et al. 2012). MESMA decomposes mixed pixels into fractions of spectrally pure endmembers (EMs). Spectral variability and the mixing complexity are accounted for by iteratively calculating multiple linear mixture models with varying number and types of EMs. MESMA relies on the availability of a spectral library with relevant EMs. In the regression case, advanced techniques such as support vector regression (SVR; Walton 2008; Esch et al. 2009), neural networks (Pu et al. 2008; Van de Voorde et al. 2009) or regression trees (Yang et al. 2003; Im et al. 2012b) have been tested in the urban context.

These techniques incorporate benefits of machine learning, i.e., they predict the relationship between variables by fitting flexible, non-parametric and nonlinear models without a priori assumptions on data distributions (Camps-Valls and Bruzzone 2009). Regression approaches rely on the availability of training pairs of spectral signatures and related land cover fractions, which can be for example derived by combining image spectra with reference information from high resolution data (Walton 2008; Esch et al. 2009) or by synthetically mixing library spectra (Okujeni et al. 2013). The variability of urban materials and their mixing complexity must be included in the training data, based on which regression models learn form. The use of synthetic training information, combined with SVR, has been demonstrated as a promising approach for mapping spectrally complex urban land cover types by means of airborne imaging spectrometer data (Okujeni et al. 2013; Okujeni et al. 2014).

Simulating spaceborne imaging spectrometer data allows both assessing the potential of upcoming hyperspectral satellite missions for urban applications and testing and enhancing quantitative mapping techniques. For example, Roberts et al. (2012) explored the utility of HyspIRI at 60 m spatial resolution for quantifying VIS extended by a non-photosynthetic vegetation component using simulated HyspIRI data and MESMA. The authors report a high potential of HyspIRI for mapping vegetation cover, however, also acknowledge difficulties in discriminating impervious and soil surfaces due to the amplified mixing at 60 m resolution. Specifically, the EnMAP satellite mission holds great potential for urban remote sensing (Heldens et al. 2011). The EnMAP concept comprises major improvements compared to existing spaceborne hyperspectral sensors (Segl et al. 2012). The ground sampling distance of 30 m at nadir marks an intermediate step between airborne systems and prospective HyspIRI data. With data acquisitions in the wavelength region between 420 – 2,450 nm in 244 bands and a nominal revisit time of 21 days, EnMAP represents a promising hyperspectral supplement to established multispectral systems with similar spatial resolutions such as Landsat. The mission is currently in the development and production phase (Phase D), with an expected launch in 2017. More detailed information on the EnMAP mission and instrument specifications are provided in Stuffler et al. (2009) and Kaufmann et al. (2008). Latest information on the mission is available at www.enmap.org.

The overarching goal of this paper is to assess the potential of EnMAP data for urban remote sensing. We quantified land cover along the urban-rural gradient of Berlin, Germany, using a simulated hyperspectral EnMAP scene at 30 m spatial resolution, an

airborne Hyperspectral Mapper (HyMap) scene at 9 m spatial resolution and a simulated multispectral Landsat ETM+ image at 30 m spatial resolution. We combined SVR with synthetically mixed training data for sub-pixel mapping (Okujeni et al. 2013). The specific objectives of this paper are (1) to explore the use of EnMAP data for VIS and extended VIS mapping, (2) to quantify the impact of spatial resolution on the quality of fraction maps when using spaceborne EnMAP instead of airborne imaging spectrometer data, and (3) to identify benefits of EnMAP data compared to imagery from multispectral spaceborne systems with similar spatial resolution. Our study helps to gain a deeper understanding on what the opportunities and limitations of forthcoming hyperspectral satellite mission such as EnMAP will be for urban remote sensing applications.

2 Study area, materials and methods

2.1 Study area

The study region covers subset of Berlin's urban-rural gradient (Figure IV-1). This area corresponds to the available hyperspectral image data and depicts a transect from the highly heterogeneous inner-city core to the more homogeneous peri-urban landscape around Berlin. A great variety of different urban structure types is captured within the city limits, including a commercial center, high- to low-density residential areas, industrial spaces, sport grounds, water bodies, green spaces and tree lined boulevards being characteristic for Berlin. The rural end of the gradient is characterized by different crop types or bare soil, grasslands, forest areas, small municipalities with low-density residential structures and individual industrial complexes. The study area depicts a highly diverse urban-rural gradient representative for the city of Berlin and many cities in general. Moreover, due to the high spectral diversity and spatial variability of artificial and natural surface materials, a spectrally and spatially challenging environment was selected for this study.

2.2 Image data

Two airborne HyMap images were acquired over the study area on 20th August 2009 around solar noon. HyMap acquires 128 bands with spectral sampling distances between 13 and 17 nm in the spectral region from 440 to 2,500 nm (Cocks et al. 1998), of which 111 were retained after removing noisy bands. Image pre-processing encompassed system

correction, radiometric correction and parametric ortho-rectification (Cocks et al. 1998; Richter and Schläpfer 2002; Schläpfer and Richter 2002).

Both HyMap images were captured at different altitudes along the same nadir line. The first HyMap scene was acquired at an altitude of 2,005 m (AGL), resulting in an area cover of 2.3 by 22.8 km and a spatial resolution of 3.6 m after pre-processing. The second HyMap scene (in the following referred to as *HyMap*) was acquired at an altitude of 4,750 m (AGL), resulting in an area cover of 5.5 by 22.8 km and a spatial resolution of 9 m after pre-processing (Figure IV-1). A simulated EnMAP scene (*EnMAP*) with 30 m spatial resolution (Figure IV-1) was derived from *HyMap* using the EnMAP end-to-end simulator (Segl et al. 2012). Simulations were constrained to the spatial domain, as the EnMAP spectral resolution is higher than that of HyMap. A simulated Landsat ETM+ image (*Landsat*) with 30 m spatial resolution and six optical bands was derived by spectrally resampling *EnMAP* with the Landsat ETM+ spectral filter function.

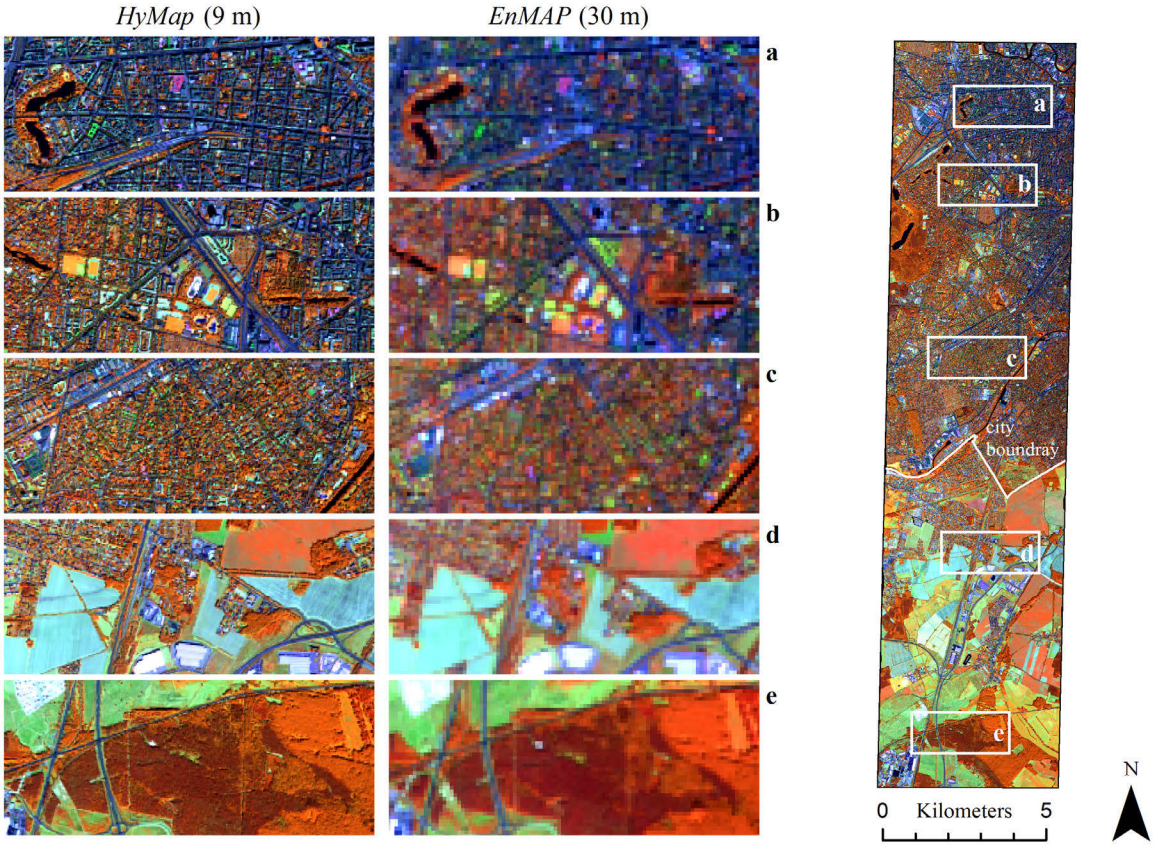


Figure IV-1: Overview of the study area and subsets of HyMap and EnMAP data showing the diversity of land cover along the urban-rural gradient. (a) high-density-, (b) medium-density-, (c) low-density urban area, (d) mixed peri-urban (arable land, industrial complexes, small municipality), and (f) forest area. All images are displayed using R = 833 nm, G = 1652 nm, B = 632 nm.

The 3.6 m HyMap image was used to develop a spectral library of urban materials (compare Section 2.3). The library, thus, constitutes an independent source for generating synthetically mixed training data (compare Section 2.4), which follows they idea of using universal data bases. *EnMAP*, *HyMap* and *Landsat* were used for fraction mapping at 9 m and 30 m spatial resolution (compare Section 2.4). We selected the 9 m image as representative for mapping from high resolution airborne data reasoning that the image was basis for simulations and thus provides radiometric and spatial consistency with the 30 m data.

2.3 Spectral library

A spectral library comprising of 75 artificial and natural surface material spectra relevant for the study region was developed for the analysis (Table IV-1). According to our objectives, the spectral library was hierarchically structured into two levels: *Impervious*, *vegetation* and *soil* (including bare soil and sand) represent urban land cover according to the VIS framework. *Roof* and *pavement* as well as *low vegetation* (including grassland, pastures and agricultural crops) and *tree* extend the VIS framework. An additional *other* category was used to assign surfaces with marginal occurrence (e.g., water, sport grounds covered with artificial turf or tartan, etc.), and did not constitute a category of interest.

We used the 41 spectra from Okujeni et al. (2013) and complemented another 34 spectra. All spectra were extracted from the 3.6 m HyMap image as mean reflectance spectra of spectrally homogeneous regions of interest. Regions for spectral sampling were identified based on field mapping data, digital orthophotos and Google Street View. Complemented spectra accounted either for additional surface components along the urban-rural gradient or for spectral variability of materials, e.g., variations due to differences in brightness or material condition. The library development followed the same iterative approach as presented in Okujeni et al. (2013).

Impervious (30 spectra), *vegetation* (31 spectra) and *soil* (4 spectra) represent broad land cover categories. Yet, even when using airborne or spaceborne imaging spectrometer data the spectral differentiation of some *impervious* and *soil* surface materials is not straightforward (Weng et al. 2008; Franke et al. 2009; Roberts et al. 2012). Extended VIS surfaces represent spectrally challenging sub-categories. *Roof* (23 spectra) is a spectrally heterogeneous category, characterized by high diversity and variability of different roofing materials. *Pavement* (7 spectra) is a category with a low number of different materials, most of them with high similarity when compared to dark roofing materials, soil mineral

matter or shaded areas. The diversity of *low vegetation* (18 spectra) is high due to variety of different grass types and agricultural crops. Likewise high spectral similarities exist between different vegetation-covered areas, which impedes the differentiation of *low vegetation* and *tree* (13 spectra).

Table IV-1: Hierarchically structured spectral library.

VIS	Extended VIS	Urban materials	No. spectra
Impervious	*Roof*	Red clay tile	4
		Red cement tile	3
		Bitumen	5
		Brown tile	1
		Black tile	1
		Brown shingle	1
		White roof material (polyethylene)	1
		White roof material (unknown)	4
		Greenish roof material (unknown)	1
		Light gray roof material (unknown)	1
		Zinc	1
	Pavement	Pavement - asphalt	4
		Pavement - concrete	3
Vegetation	*Low vegetation*	Grass (intensively manicured)	2
		Grass (extensively manicured)	1
		Grass (agricultural grassland)	3
		Grass (meadow clover)	2
		Grass (meadow alfalfa, lucerne)	1
		Grass (dry)	2
		Grass (dry agricultural grassland)	1
		Sunflower	1
		Corn	1
		Sorghum	1
		Sugarbeet	1
		Potatoes	2
	Tree	Deciduous tree	11
		Coniferous tree	2
Soil		Bare soil	2
		Sand (playground)	2
Other		Red sand (cinder field)	1
		Red sand (cinder court)	2
		Railtrack	2
		Artificial turf	2
		Tartan (sports ground)	1
		Water	2

Multiple spectra per material considered to account for spectral variability due to variations in illumination and shading effects or variations in material condition.

2.4 Support vector regression modeling using synthetically mixed training data

Empirical regression modeling relies on the availability of continuous training information, i.e., pairs of spectral signatures and related land cover fractions. In this study, we combined SVR with synthetically mixed training data for sub-pixel mapping (Okujeni et al. 2013). Synthetically mixed data are sets of pure material spectra, multiple mixed spectra thereof as well as associated mixing fractions. Such sets are generated from a spectral library, provide valuable training information for regression modeling of single target land cover categories and enable utilizing the strength of machine learning-based SVR to cope with spectral complex urban land cover types, particularly when quantitative training data is missing (Okujeni et al. 2013; Okujeni et al. 2014).

We followed the description in Okujeni et al. (2013), which included (i) the partitioning of the spectral library into a target (e.g., *impervious*) and a background category, which includes all remaining categories (e.g., *vegetation*, *soil*, *other*), (ii) the calculation of synthetic mixtures between pure target category spectra with 100% mixing fraction and pure background category spectra with 0% mixing fraction of the respective target category, (iii) the training of a SVR model, and (iv) the derivation of a fraction map of the respective target category with subsequent post-processing.

The generation of synthetically mixed training data assumes a linear mixing process between material spectra, an assumption regularly made in the context of urban mapping (Small and Lu 2006; Powell et al. 2007; Franke et al. 2009; Roberts et al. 2012). To account for the mixing complexity (i.e., number of materials that contribute to the synthetic mixture), we included binary and selected ternary linear mixtures of materials. Binary mixtures comprised all possible pairs between target and background category spectra. Ternary mixtures comprised typical combinations between three materials, including mixtures of a target category spectrum and two background spectra or vice versa. We thereby considered both combinations between different categories (e.g., “deciduous tree - red clay tile - asphalt” or “deciduous tree - grass - soil”) and combinations within categories (“deciduous tree - red clay tile - bitumen” or “asphalt - concrete - tree”). Knowledge on relevant ternary mixtures for the study area was derived from visual inspection of digital orthophotos. The selective approach enabled the inclusion of higher order mixtures and avoided an excessive amount of training samples from operative use of all possible ternary material combinations. The mixing interval (i.e., number of intermediate mixtures within the 0-100% mixing range) was set to 20%, resulting

increments of 20%, 40%, 60% and 80% when linearly mixing target against background category spectra. This interval has been shown to provide sufficient continuous training information for SVR interpolations (Okujeni et al. 2013; Okujeni et al. 2014). Synthetically mixed training data were generated for VIS and extended VIS target categories *impervious*, *vegetation* and *soil*, and *roof*, *pavement*, *low vegetation* and *tree*, respectively. The *other* category was always used in the background.

The synthetically mixed training data sets were used as input for SVR modeling. SVR emanates from the field of machine learning as flexible, non-parametric and nonlinear quantitative modeling technique. Detailed introductions into underlying concepts of SVR were provided by Vapnik (1995) or Schölkopf and Smola (2002). SVR fits a hyperplane to a set of training data to predict the dependency between n-dimensional input features (i.e., spectral bands) and a 1-dimensional target variable (i.e., fractions of a target category). The optimal approximating hyperplane is found by quadratic optimization. Nonlinear problems are solved by making use of a kernel function. SVR modeling was carried out with imageSVM (Rabe et al. 2014b), which is an open source tool delivered as part of the EnMAP-Box (Rabe et al. 2014a). The SVR model training required the selection of parameters for the Gaussian kernel function, the epsilon-insensitive loss function and for regularization. This was done by means of a fully automated parameter tuning via grid search. Subsequently, SVR models were applied to *EnMAP*, *HyMap* and *Landsat* to derive land cover fraction maps at 9 m and 30 m.

SVR mappings covered continuous fractions between 0 and 100% as well as negative (< 0%) and super-positive (> 100%) values. This estimation of physically unrealistic fractions relates to improper extrapolation (Okujeni et al. 2013) and was compensated for by setting negative and super-positive fractions to 0% and 100%, respectively.

2.5 Reference data and validation

We performed a polygon-wise statistical evaluation of results using urban blocks from the Urban and Environmental Information System of Berlin (SenStadt 2014b). Polygon-wise accuracy assessment is a common strategy to evaluate the quality of fraction maps (Powell et al. 2007; Roberts et al. 2012), whilst mitigating the influence of inaccurate co-registration between estimate and reference (Stehman and Wickham 2011). Beyond, urban blocks constitute a suitable spatial unit for urban management or urban environmental studies (Pauleit and Duhme 2000; Heiden et al. 2012).

We used 74 polygons, including 56 building blocks of different densities, 6 street polygons and 12 green spaces. All polygons were selected within a ±10° off-nadir region to minimize distortions due to reflectance anisotropy (Schiefer et al. 2006) or urban 3D-geometry (van der Linden and Hostert 2009). The urban blocks included a wide spectrum of different anthropogenic and natural cover types. We added 6 polygons to sufficiently represent the fraction range of soil surfaces. Polygon boundaries were adjusted to exclude image pixels that only partially fell within reference polygons (Figure IV-2).

The area estimate within the 80 polygons was based on a high resolution reference map. This reference map was created on the basis of digital cadastral information (SenStadt 2014a), a digital surface model and digital orthophotos. Validation of all fraction maps was carried out by comparing estimated and reference fractions. We calculated the mean absolute error (MAE) as measure of accuracy. To evaluate the goodness of fit, we further used the coefficient of determination (R^2), as well as the slope and the intercept of a fitted least squares linear regression model.

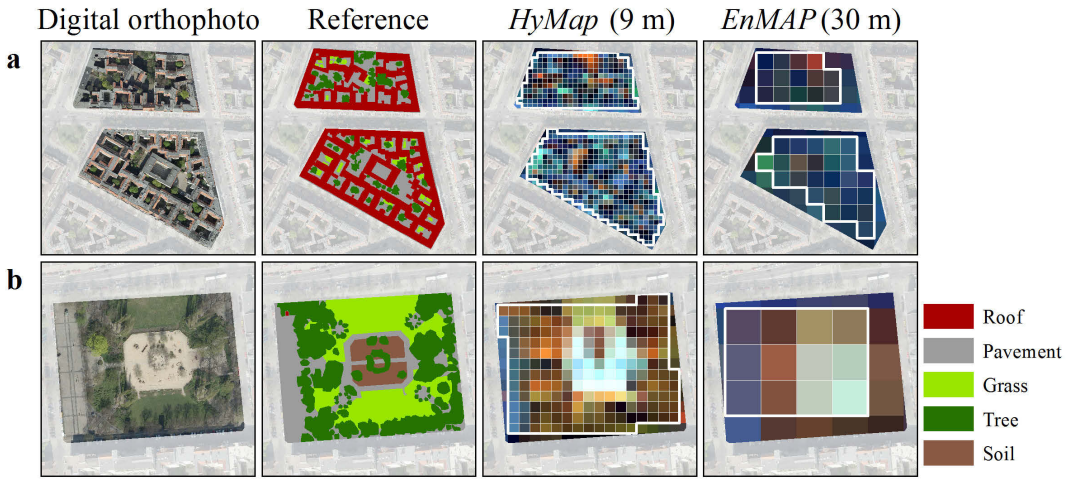


Figure IV-2: Examples of urban blocks and polygons used for validation. (a) high-density urban area, and (b) green spaces.

3 Results

3.1 Mapping VIS fractions along the urban-rural gradient

The comparison between reference and estimated fractions at urban block scale provides a statistical evaluation of the quality of VIS fraction maps (Figure IV-3). *Impervious* (MAE = 10.9%, R^2 = 0.87), *vegetation* (MAE = 10.5%, R^2 = 0.85) and *soil* (MAE = 11.9%, R^2 = 0.87) estimates from *EnMAP* show strong linear relationships close to the 1:1 line when

compared to reference data. A general overestimation of low fractions and underestimation of high fractions is observed for all land cover types. Most obvious prediction errors include overrated *impervious* fractions within green spaces and overrated *soil* fractions within urban blocks where bare soil surfaces are inexistent.

Comparisons between accuracies of *EnMAP* and *HyMap* reveal a rather small impact of the coarser spatial resolution on the quality of fraction maps. *Impervious*, *vegetation* and *soil* estimates from *HyMap* are slightly more (or even similarly) accurate. Comparisons between accuracies of *EnMAP* and *Landsat* reveal an important difference when using multispectral information. While accuracies of *vegetation* and *soil* appear very similar, *impervious* estimates derived from *Landsat* are significantly less accurate (MAE = 15.5%, $\Delta R^2 = 0.77$). Particularly *impervious* fractions within green spaces and urban blocks from the low-density urban area are overrated in *Landsat*.

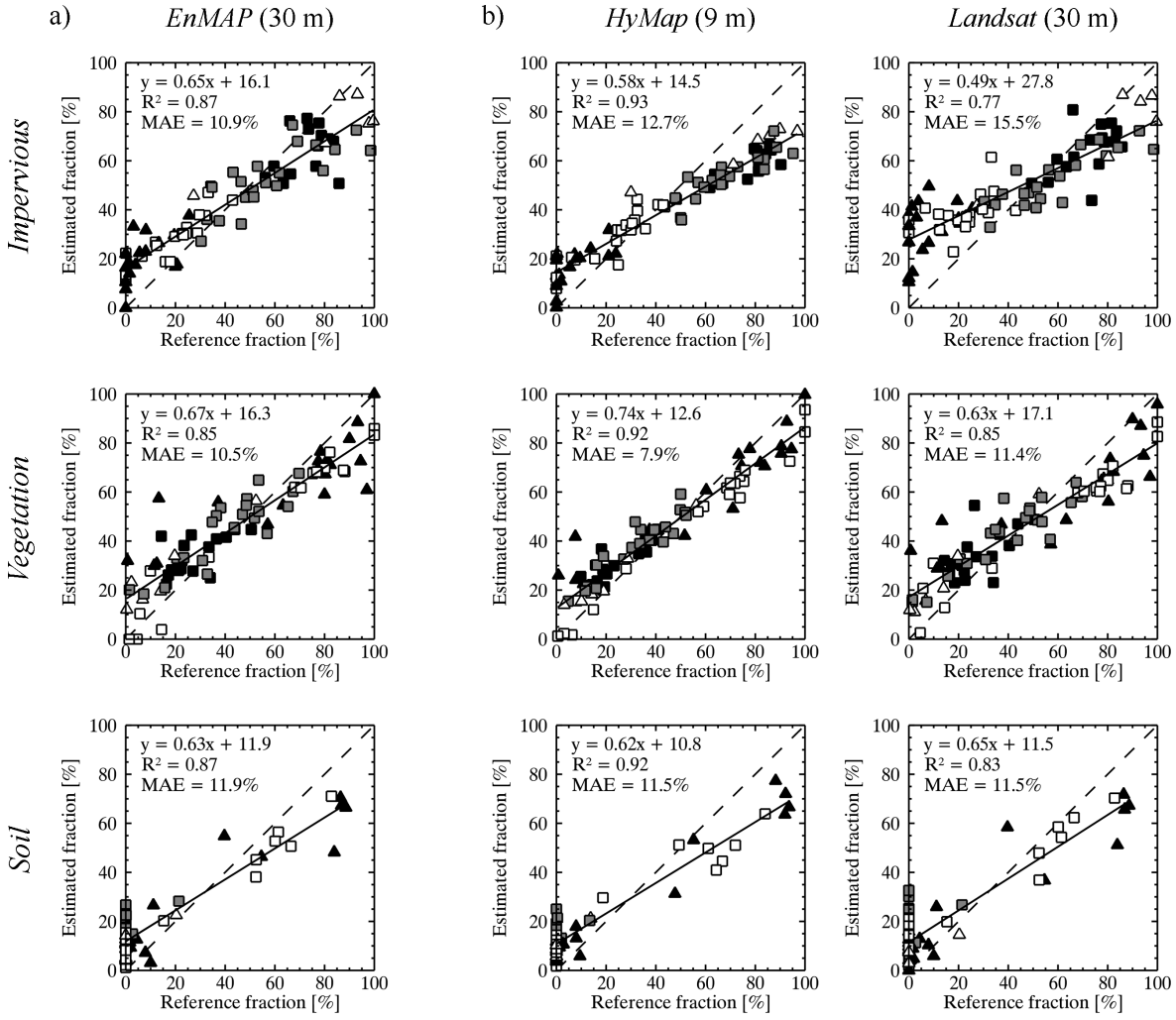


Figure IV-3: Scatterplots of *impervious*, *vegetation* and *soil* fractions from the reference data compared to modeled fractions derived from a) *EnMAP*, and b) *HyMap* and *Landsat*. Square symbols = building blocks (black = high-density-, gray = medium-density-, white = low-density urban area); triangle symbols = green spaces (black), street polygons (white).

Fraction maps in Figure IV-4 depict the land cover composition along Berlin's urban-rural gradient well according to the VIS model. Mappings derived from *EnMAP* visualize high densities of *impervious* surface types (displayed in red) within the city core. Apart from few peri-urban municipalities and industrial zones, this density decreases towards the urban fringe. *Vegetation* cover (displayed in green) reveals the opposite trend with gradually increasing fractions towards the urban fringe and the countryside. *Soil* surface types (displayed in blue) marginally occur within the city, with higher fractions mainly found on arable land. Low confusion between the spectrally similar *impervious* and *soil* surface cover types is observed in *EnMAP*. *Other* surface types were only used as background category during processing and thus appear dark in the RGB composite.

Despite the higher level of spatial detail of the VIS fraction maps at 9 m, patterns of fraction distributions between *EnMAP* and *HyMap* are consistent. In contrast, fraction maps derived from *Landsat* reveal obvious differences compared to maps derived from *EnMAP*. *Impervious* estimates are generally higher, which appears more subtle within the dense urban core and more obvious towards the urban fringe. Confusion between *impervious* and *soil* surfaces as well as *impervious* and *vegetation* is found in the mixed peri-urban zone and the forest area.

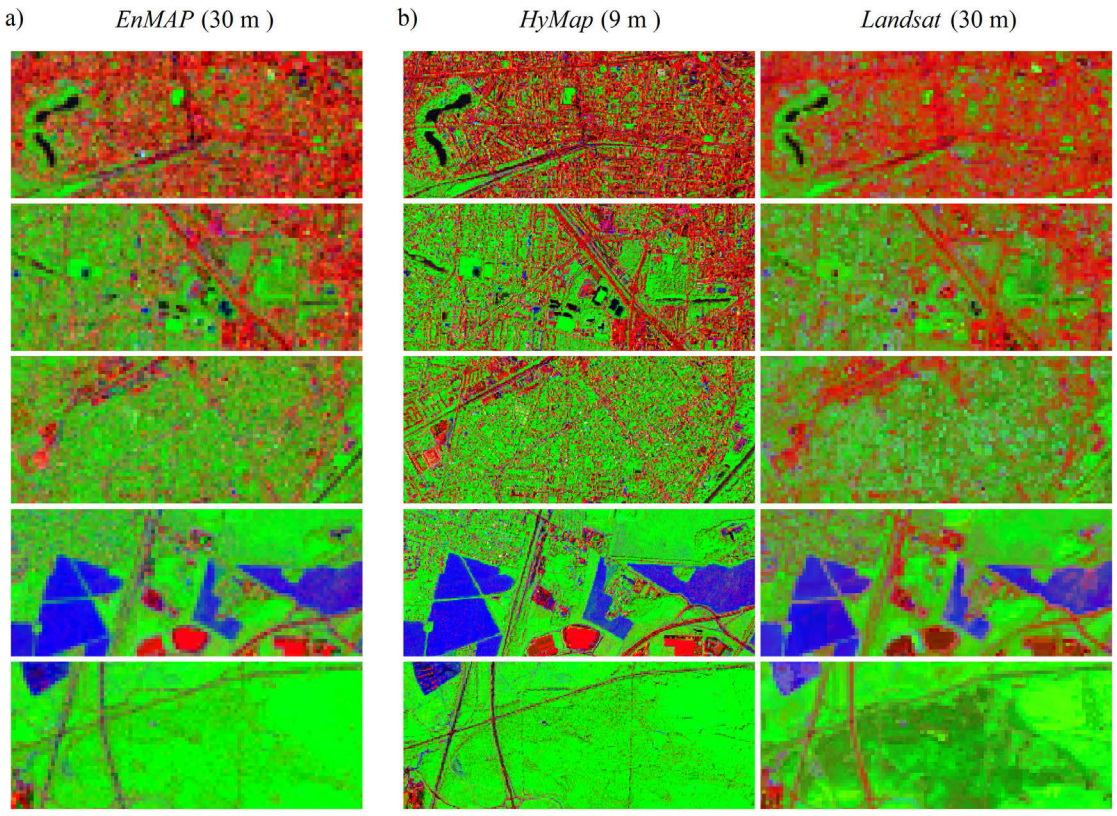


Figure IV-4: *Impervious* (red), *vegetation* (green) and *soil* (blue) fraction maps derived from (a) *EnMAP*, and (b) *HyMap* and *Landsat*.

3.2 Extending VIS mapping along the urban-rural gradient

The statistical evaluation of extended VIS land cover maps is illustrated in Figure IV-5. *Roof* (MAE = 12.4%, R^2 = 0.52), *pavement* (MAE = 21.0%, R^2 = 0.68), *low vegetation* (MAE = 7.3%, R^2 = 0.84) and *tree* (MAE = 16.6%, R^2 = 0.70) estimates from *EnMAP* show close to linear relationships along the 1:1 line when compared to the reference data. *Roof*, *pavement* and *tree* estimates reveal a considerable scatter. Major prediction errors include mapped *roof* fractions within green spaces and street polygons, overrated *pavement* estimates at small to intermediate fractions and a nearly constant overestimation of *tree* fractions across the fraction range.

Fraction estimates derived from *HyMap* are generally more accurate compared to results from *EnMAP*. However, the impact of the coarser spatial resolution on the quality of maps mainly affects *roof* estimates (MAE = 10.7%, R^2 = 0.74), while mapping of *pavement, low vegetation* and *tree* only improve slightly. Fraction maps derived from *Landsat* are less accurate compared to results from *EnMAP*. Particularly low to intermediate *roof* fractions are highly overestimated and *low vegetation* estimates deviate from the 1:1 line with a considerable increase in scatter.

Fraction maps of *roof* (displayed in red), *low vegetation* (displayed in blue) and *tree* (displayed in green) (Figure IV-6) provide a considerably more detailed insight into the land cover composition along the urban-rural gradient than VIS. *Pavement*, *soil* and *other* surface types appear dark in the RGB composite. *Roof* fractions gradually decrease from the city center towards the urban fringe. Rooftops of large industrial buildings in the peri-urban surrounding are well captured. Vegetation-covered areas within the city are dominated by *tree*. Except for larger lawns in the city (e.g., sport grounds, parks), most *low vegetation* types concentrate in the peri-urban area.

The comparison between *EnMAP*- and *HyMap*-based fraction maps reveals overall consistent spatial patterns in land cover fraction distribution. In contrast, *Landsat*-based fraction maps show much more confusion between the different land cover types. The spatial pattern of *roof* rather follows the distribution observed for the broad *impervious* category (Figure IV-4). Considerable confusion between *roof* and *soil* and *low vegetation* and *tree* is particularly observed in the mixed peri-urban zone and the forest area.

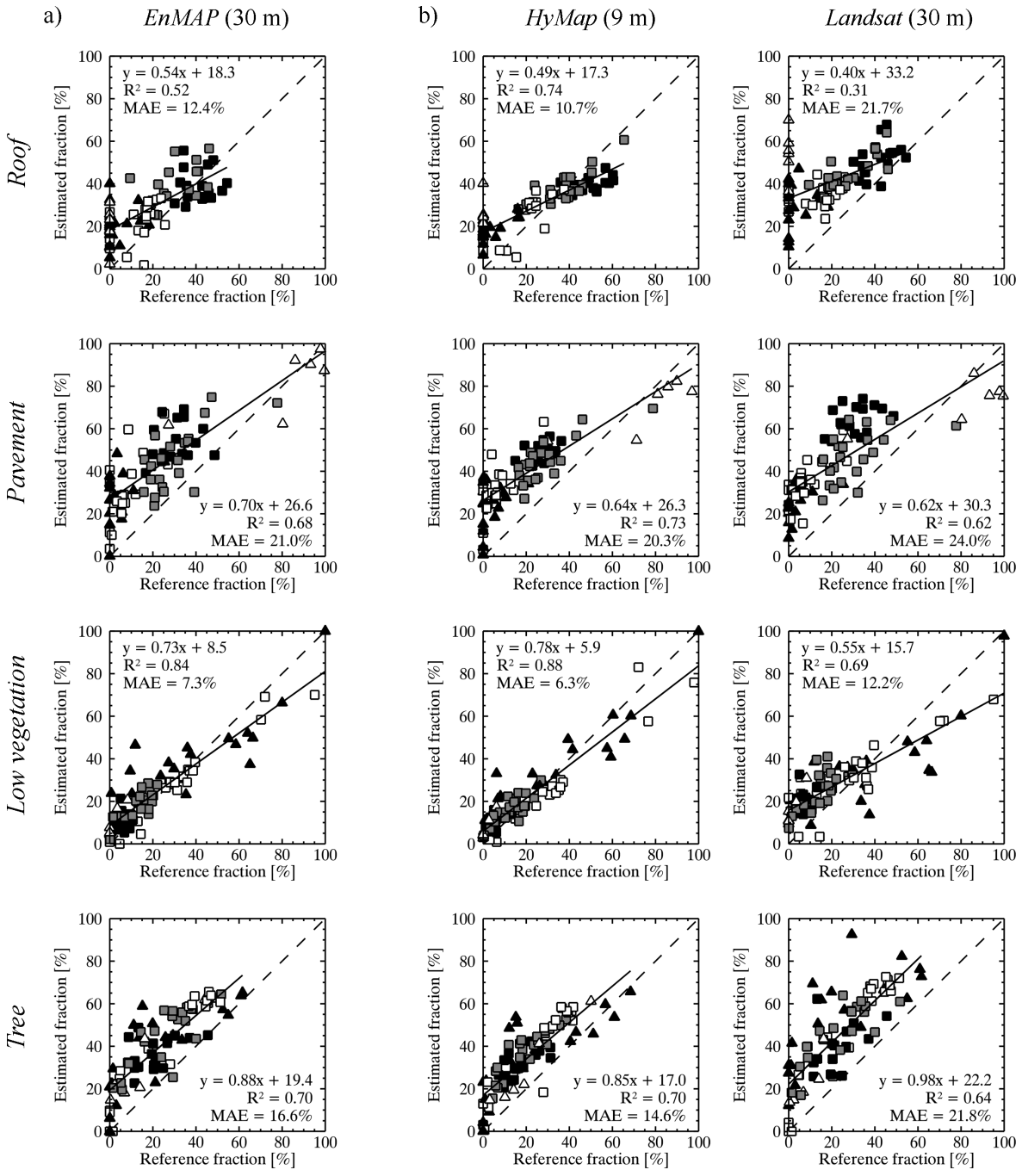


Figure IV-5: Scatterplots of *roof*, *pavement*, *low vegetation* and *tree* fractions from the reference data compared to modeled fractions derived from a) *EnMAP*, and b) *HyMap* and *Landsat*. Square symbols = building blocks (black = high-density-, gray = medium-density-, white = low-density urban area); triangle symbols = green spaces (black), street polygons (white).

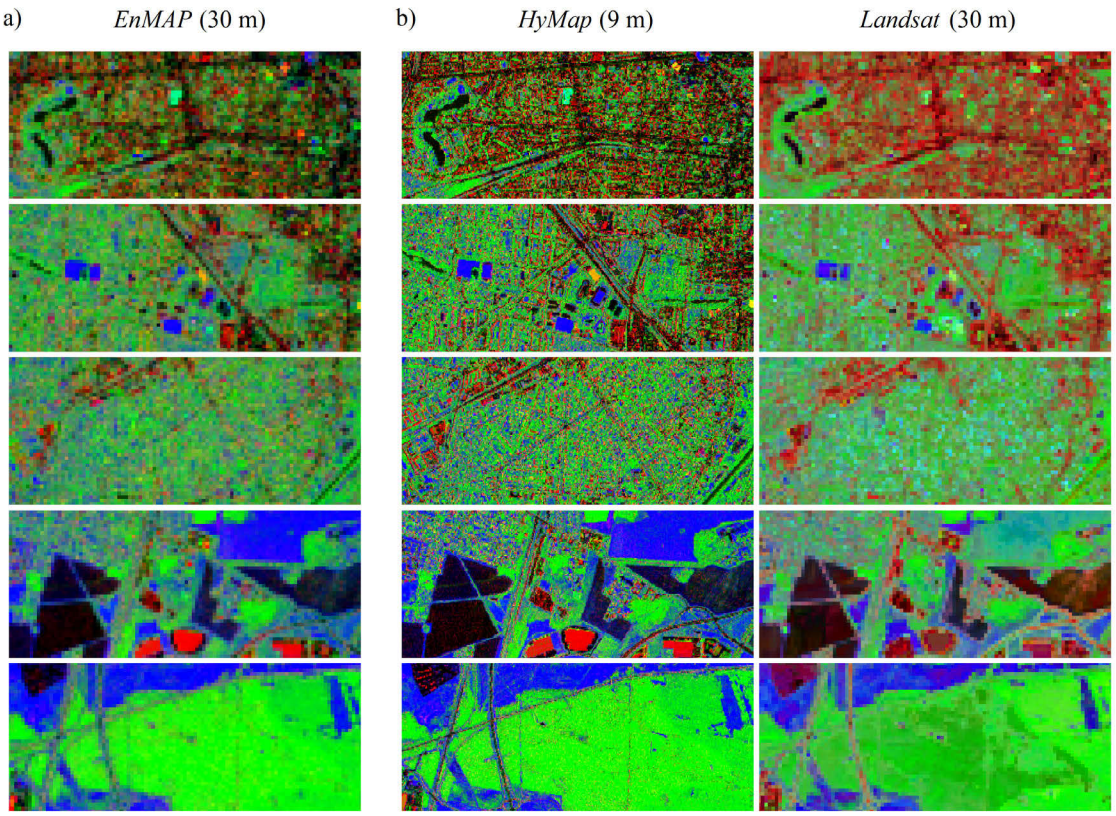


Figure IV-6: *Roof* (red), *tree* (green) and *low vegetation* (blue) fraction maps derived from (a) *EnMAP*, and (b) *HyMap* and *Landsat*.

4 Discussion

Data acquired by future spaceborne hyperspectral imagers hold great potential for mapping challenging heterogeneous environments. Based on simulated data, we assessed the potential of the hyperspectral satellite mission EnMAP for urban land cover assessments on a purely spectral basis. This is a suitable strategy for investigating opportunities and limitations of upcoming remote sensing systems for environmental studies (Guanter et al. 2009; Roberts et al. 2012; Laurent et al. 2014). The simulation of *EnMAP* was constrained to the spatial domain. This was due to the higher spectral resolution of the EnMAP sensor compared to the HyMap sensor. Yet, our investigations considered a major challenge when adopting spatially simulated EnMAP data at 30 m resolution, namely the mixed pixel problem that is expected to especially complicate studies in spatially and spectrally heterogeneous urban environments.

Our first objective was to explore the use of EnMAP data for VIS and extended VIS mapping. By using *EnMAP*, we were able to show that EnMAP acquisitions will be well suited for producing accurate fraction maps of the broader *impervious*, *vegetation* and *soil*

categories along the urban-rural gradient (see Figures IV-3 and IV-4). VIS mapping accuracies are high, with MAE values below 12.0% and R^2 values of 0.85 and greater. Despite different methodological approaches, data characteristics, urban setups and validation units, this accuracy points to possible advantages of EnMAP data characteristics compared to existing and upcoming spaceborne imaging spectrometers for urban applications: Roberts et al. (2012) reported R^2 values between 0.41 and 0.81 for VIS mapping using simulated HyspIRI data at 60 m, and Weng et al. (2008) reported a MAE of 14.8% and a R^2 of 0.73 for mapping *impervious* cover using Hyperion data at 30 m. Observed mapping uncertainties are small and result from known spectral similarities that are difficult to overcome completely, even when using hyperspectral information. This includes the spectral confusion between *impervious* and *soil* materials (Weng et al. 2008; Roberts et al. 2012) as well as between dark *impervious* materials and shaded surfaces within stands of vegetation (Okujeni et al. 2013). The extended VIS mapping differentiates *impervious* and *vegetation* into the sub-categories *roof* and *pavement* as well as *low vegetation* and *tree*. By using *EnMAP*, we showed that EnMAP data will allow extending VIS by such detailed land cover types (see Figures IV-5 and IV-6). However, accuracies are lower (MAE between 7.3% and 21.0%, R^2 between 0.52 and 0.84) compared to the broader VIS mapping, which can be similarly expected for EnMAP-based mapping. Particularly the precise description of *roof* and *pavement* fractions at 30 m will be challenging. Observed overestimations relate to spectrally similar materials (e.g., bitumen roof vs. street asphalt) or – as mentioned previously – the confusion between dark anthropogenic materials and shaded surfaces within stands of vegetation. Similar findings were reported for studies using higher resolution airborne data (Herold et al. 2004; van der Linden et al. 2007; Franke et al. 2009; Okujeni et al. 2013). *Low vegetation* and *tree* mapping from *EnMAP* are more precise and reveal less confusion indicating that EnMAP data will be well suited for a better differentiation of vegetation cover along urban-rural gradients. The systematic overestimation of *tree* fractions may be in part explained by highly complex illumination and shading effects between different deciduous and coniferous tree stands. The general problem of spectral similarity and shaded areas when mapping detailed urban land cover types is highly complex and challenging when using global regression models for mapping. Possible solutions for these shortcomings go beyond the scope of this work and may include brightness normalization (Wu 2004; Feilhauer et al. 2010) or spectral feature extraction (Herold et al. 2004; Heiden et al. 2007) prior to SVR modeling.

Our second objective was to quantify the impact of spatial resolution on the quality of fraction maps when using EnMAP instead of higher resolution airborne imaging spectrometer data. This change in spatial scale implies a loss in spatial detail and a significant increase in the number of mixed pixels. Results from *EnMAP* at 30 m and *HyMap* at 9 m generally reveal reduced accuracies for mapping assessments on the coarser resolution data. This can be similarly expected for EnMAP-based mapping and relates to the increasingly difficult discrimination of materials at 30 m resolution, especially those with high spectral similarities or low spectral contrast. However, the impact of amplified mixing on VIS mapping quality is rather low (see Figures IV-3 and IV-4) and *impervious*, *vegetation* and *soil* estimates appear robust in terms of spatial scale (max. net decrease in accuracies: ΔMAE = +2.6%, ΔR^2 = -0.07). For extended VIS results (see Figures IV-5 and IV-6), the *roof* mapping derived from *EnMAP* appears more prone to the change in spatial scale (net decrease in accuracy: ΔMAE = +1.7%, ΔR^2= -0.22), while the impact of spatial resolution on *pavement*, *low vegetation* and *tree* mapping is rather low (max. net decrease in accuracies: ΔMAE = +2.0%, ΔR^2 = -0.05). The mapping of heterogeneous urban surface types with high spectral diversity and variability at sub-pixel scales will therefore require special attention, as exemplified with the reduced accuracies for *roof* mapping. This generally points to the advantage of higher spatial resolution airborne images over coarser resolution spaceborne data for urban applications. Yet, airborne acquisitions are limited in synoptic coverage at urban scales, which impairs their use for mapping urban land cover globally.

Our third objective was to identify benefits of EnMAP data compared to imagery from multispectral spaceborne systems with similar spatial resolution. In most cases, fraction maps at 30 m proved to be more accurate when using hyperspectral *EnMAP* instead of multispectral *Landsat* (see Figures IV-3, IV-4, IV-5 and IV-6). Our study particularly highlights the superiority to improve *impervious* mapping when adopting the VIS scheme (net increase in accuracy: ΔMAE = -4.6%, ΔR^2= +0.1) and the possibility to extend VIS by more detailed urban surface types (net increase in accuracies: ΔMAE range from -3.0% to -9.3%, ΔR^2 range from +0.06 to +0.21). This can be attributed to the presence of more distinct material-specific absorption features provided by the continuous spectral information that outweighs the mixed pixel problem at coarser scales. As simulating EnMAP's full spectral characteristics from *HyMap* was not possible, we assume that our results represent conservative estimates. Mapping accuracies will hence rather underestimate the full potential of EnMAP and its improvement in urban mapping

compared to approaches based on data from Landsat-like sensors. Advantages of hyperspectral information content from airborne data for material-oriented urban mapping (Herold et al. 2003; Gamba and Dell'Acqua 2006) and spaceborne Hyperion data for *impervious* mapping (Weng et al. 2008) have already been demonstrated. Here, we demonstrate the expectable surplus of data acquired by future spaceborne hyperspectral missions, such as EnMAP, over data delivered by established multispectral satellite systems, such as Landsat, for improved and extended VIS urban land cover assessments on a purely spectral basis.

The combination of support vector regression with synthetically mixed training data proved to be a suitable approach for sub-pixel mapping of land cover along the urban rural-gradient. Spectrally heterogeneous urban categories (e.g., *roof*) and spectrally similar urban land cover types (e.g., *impervious* vs. *soil*, *roof* vs. *pavement* or *low vegetation* vs. *tree*) were well quantified. This finding is in accordance with a previous study that demonstrated the similar or higher performance of the approach compared to traditional MESMA (Okujeni et al. 2013). The strength of SVR to model the complex functional relation between multiple mixtures and associated mixing fractions in a single global model per target category was exploited. Spectral diversity and variability of urban surface types was included in the synthetic mixtures and learned during model training. The strategy to consider only binary and selected ternary mixtures avoided an excessive amount of training samples and high processing times. Possible other ternary combinations or even higher order mixtures were not accounted for during regression modeling. This may have affected the quality of mappings, particularly for 30 m resolution *EnMAP* and *Landsat*. Yet, results of this study are valid, which points to the capability of the SVR models to account for large parts of the spectral mixing within both higher resolution airborne and coarser resolution spaceborne imagery. By mapping land cover along the urban-rural gradient, the approach's capability to accurately describe transition zones between highly heterogeneous urban areas with amplified mixed pixels and more homogenous peri-urban landscapes with less mixing was demonstrated.

5 Conclusion

The hyperspectral satellite mission EnMAP is designed to provide timely and frequent imaging spectrometer data globally. In this paper, we assessed the potential of such imagery for urban remote sensing. We used simulated data (1) to explore the use of

EnMAP data for VIS and extended VIS mapping, (2) to quantify the impact of spatial resolution on the quality of fraction maps when using spaceborne EnMAP instead of airborne imaging spectrometer data, and (3) to identify benefits of EnMAP data compared to imagery from multispectral spaceborne systems with similar spatial resolution. The study area of Berlin, Germany, offered the opportunity to quantitatively assess opportunities and limitations along a representative, spatially and spectrally heterogeneous urban-rural gradient.

We conclude that EnMAP data will open up new opportunities for both improved and extended VIS urban land cover assessments. SVR and synthetically mixed training data proved to be a suitable technique for sub-pixel mapping. This suggests that the combination of spaceborne imaging spectrometer data and machine learning holds great potential for introducing new quality into urban remote sensing. EnMAP will significantly increase the availability of imaging spectrometer data, yet bears the drawback of a larger number of spectrally mixed pixels compared to spatially constrained airborne acquisitions. This will particularly challenge the mapping of spectrally complex, heterogeneous urban land cover types. At the same time, the hyperspectral information of 30 m pixels will provide a considerable greater material discrimination capability. This will be a major advantage of EnMAP for urban remote sensing, particularly compared to established multispectral systems with similar spatial resolutions such as Landsat.

The upcoming EnMAP mission will allow moving from individual case studies to global urban mapping assessments by means of hyperspectral data. Follow-up studies on imagery from cities around the world are necessary to assess the full potential of hyperspectral satellite data. Additional work should focus on generating a universal urban spectral library representative for many urban areas, which can be used to make quantitative mapping techniques such as the presented SVR approach transferable. Investigations should consider the temporal domain in order to explore benefits of EnMAP data that may arise from the multitemporal data acquisition. Particularly phenology (Tigges et al. 2013) may contribute to an improved characterization of urban vegetation types or condition. Once the full spectral, spatial and temporal potential of EnMAP data from urban areas is exploited, a significant contribution to interdisciplinary urban environmental research and ultimately urban management can be made.

Acknowledgements

The authors thank the German Aerospace Centre for data acquisition and pre-processing during the HyEurope 2009 Campaign, Karl Segl and Christian Rogaß from the German Research Center for Geosciences for data simulations and the Berlin Senate Department for Urban Development for providing data from the Urban and Environmental Information System. This research was funded by the German Research Foundation (Project no. HO 2568/2-2) and supports the EnMAP Core Science Team activities (FKZ 50EE0949) funded by the German Federal Ministry of Economics and Technology. It further contributes to comparing multi- and hyperspectral image analysis potentials in the frame of the USGS-NASA Landsat Science Team.

Chapter V:
Synthesis

1 Summary

Remote sensing is a valuable tool to derive spatial and temporal information about the physical composition of urban areas in a systematic manner. The high information content of imaging spectrometer data in particular holds great potential for describing detailed urban surface properties on a purely spectral basis (Herold et al. 2003; van der Linden et al. 2007; Franke et al. 2009). Such descriptions are essential for developing a better understanding of the integrated human-natural system in times of rapid urbanization (Alberti 2005; Cadenasso et al. 2007). The forthcoming hyperspectral satellite missions will provide high quality imaging spectrometer data on a global scale. This may contribute to the development of new urban land cover products. However, the difference between spatially high resolved airborne and coarser resolved spaceborne data will constitute a great challenge for urban mapping. Developing universal, quantitative processing workflows and testing opportunities and limitations based on simulated data helps to gain insights into the potential of spaceborne imaging spectrometry for urban remote sensing.

The scientific focus of this work was to quantify urban land cover by means of machine learning and imaging spectrometer data at multiple spatial scales. In the first part of this work, a quantitative mapping framework was introduced and tested with respect to the method's capability to map urban land cover, its performance compared to other sub-pixel mapping techniques, and its universality in terms of adaptability and transferability when proceeding from the scale of airborne to spaceborne data. The second part of this work made a value estimation of data that will be acquired by the hyperspectral satellite mission EnMAP for urban mapping. Emphasis was put on the possibility to map and extend the universal VIS scheme and on the illustration of opportunities and limitations of EnMAP data in the urban context. The study region along the urban-rural gradient of Berlin, Germany, proved to be a representative, spectrally, and spatially challenging environment for the experiments. A great variety of urban structure types was captured, with a high diversity and variability of anthropogenic and natural materials. The spectral library of relevant urban surface materials, the airborne hyperspectral data at different spatial resolutions, and the simulated EnMAP scene constituted a rich data set for experiments. The independent reference map and auxiliary data sources, e.g., digital orthophotos or Google Street View, allowed for reliable validation of results. This helped to evaluate the

quality of mappings by identification of error sources and general limitations of remote sensing data for urban mapping on a purely spectral basis.

In the following, findings of the core research chapters (Chapter II-IV) are summarized with respect to the research questions stated in Chapter I:

Research question I: Does the combination of SVR with synthetically mixed training data allow the reliable quantification of urban land cover from imaging spectrometer data at multiple spatial scales?

In Chapter II, the concept of combining SVR with synthetically mixed training data for sub-pixel mapping was introduced. Analyses were carried out on 3.6 m HyMap data, focusing on urban zones of high-density, medium-density and low-density. Results demonstrate the suitability of the method for reliably quantifying *roof*, *pavement, grass*, and *tree* cover. Moreover, the method improved fraction estimates compared to traditional MESMA. In Chapter III, a more generic evaluation of the SVR approach was provided by comparisons to other advanced regression techniques such as artificial neural networks or random forest. Furthermore, synthetic training mixtures of higher complexity were generated to test to what extent regression models can be adapted to the amplified spectral mixing of coarser spatial resolution data. Experiments were carried out on 3.6 m and 9 m HyMap data, again focusing on the high- to low-density urban zones and considering *roof*, *pavement*, *grass*, and *tree* cover. Results reveal the superiority of kernel-based techniques in general and SVR in particular in comprehensively quantifying urban land cover and coping with coarser resolution data when trained with more complex training mixtures. In Chapter IV, SVR models were applied to simulated EnMAP data at 30 m resolution, focusing on VIS and extended VIS mapping along the urban-rural gradient. Results demonstrate the suitability of the approach to cope with coarser spatial resolution spaceborne data and to describe the transition zone between a heterogeneous urban core with large numbers of spectrally mixed pixels and a homogenous peri-urban landscape where spectral mixing is less pronounced.

The workflow of combining SVR with synthetically mixed data separates regression model training – based on synthetic data – from model application and independent validation, and thus implements the idea of a universal mapping framework. The use of synthetic mixtures is novel and poses a straightforward and flexible strategy to generate training information needed for sub-pixel mapping by means of empirical regression. The generation of synthetic mixtures requires a spectral library, which accounts for spectral

diversity and material variability within the study region. To adapt regression models to coarser resolution data, synthetic mixtures of higher mixing complexity must be included in the training data. Based on synthetic training mixtures, SVR was capable of delineating urban surface types with high within-class variability (e.g., *impervious*, *vegetation*, *roof* or *low vegetation*) and high between-class similarity (e.g., *impervious* vs. *soil*, *roof* vs. *pavement* or *grass* vs. *tree*). SVR was capable of coping with spectral mixing within both higher spatial resolution airborne and coarser spatial resolution spaceborne data. This underlines the universality of the approach and the strength of SVR to effectively utilize training data of higher mixing complexity when models are transferred across spatial scales. The superiority of SVR compared to other quantitative techniques demonstrates the general strength of kernel-based methods from the field of machine learning for land cover mapping in spectrally and spatially complex urban environments. Compared to per-pixel classification approaches, such as the support vector classifier, SVR combined with synthetically mixed training data accounts for the mixed pixel problem and thus preserves the heterogeneity of land cover.

Overall, the combination of SVR with synthetically mixed training data allowed the reliable quantification of urban land cover from imaging spectrometer data at multiple spatial scales. This underscores the universality of the approach for exploiting prospective spaceborne imaging spectrometer data. The method was tested in a challenging urban landscape and thus may also be well suited for other environments, including heterogeneous agricultural areas, mixed forests, or shrubland transition zones.

Research question II: What type of urban land cover information can be extracted from EnMAP data, and how reliable are the mapping results?

In Chapter IV, the value of data acquired by the upcoming hyperspectral satellite mission EnMAP for quantifying land cover along an urban-rural gradient was explored. A simulated EnMAP scene with 30 m resolution was derived from HyMap data. Image simulation was constrained to the spatial domain, as EnMAP's prospective spectral resolution (~244 bands) is higher than that of HyMap (128 bands). Yet, experiments considered the amplified mixed pixel scenario, which is expected to be a major challenge when adopting such imagery for urban mapping. Results demonstrate that EnMAP data will be well-suited for mapping *impervious*, *vegetation*, and *soil* surfaces according to the VIS framework, and for extending VIS mapping with more detailed sub-categories such as *roof*, *pavement*, *low vegetation*, and *tree*. A comprehensive analysis and discussion of the

benefits and limitations of EnMAP data compared to hyperspectral airborne and multispectral spaceborne imagery was additionally provided. These findings suggest that the coarser spatial resolution of EnMAP will generally lead to reduced mapping accuracies in heterogeneous urban environments compared to results from high spatial resolution airborne data. However, the impact of the larger numbers of mixed pixels will be rather low, unless spectrally challenging urban land cover types (e.g., *roof*) are delineated. In contrast, EnMAP acquisitions will significantly improve urban land cover assessments compared to results from multispectral data with similar spatial resolution such as Landsat. Results highlight the superiority of EnMAP data for improving *impervious* estimates and for facilitating the mapping of more detailed surface types along the urban-rural gradient.

Findings underline the value of EnMAP data for urban mapping and thus EnMAP's potential to derive new urban land cover products. The quality of VIS mapping was high (MAE between 10% and 12%, R^2 between 0.85 and 0.9), with minor confusion caused by the spectral similarity of some impervious-, soil-, and shaded surface covers. The quality of extended VIS mapping was lower, with satisfactory results for *roof* and *pavement* (MAE of 12.4% and 21.0%, R^2 of 0.52 and 0.68) and accurate results for *grass* and *tree* (MAE of 7.3% and 16.6%, R^2 of 0.84 and 0.7). Confusion caused by spectrally similar urban materials when differentiating impervious surface types and the influence of shade were reported as major sources of error. Such errors were already found in studies using airborne hyperspectral data (Herold et al. 2003; van der Linden et al. 2007; Franke et al. 2009), and they highlight general limitations of urban mapping on a purely spectral basis.

Overall, results suggest that EnMAP data will allow reliable quantifications of the broad urban land cover categories *impervious*, *vegetation*, and *soil*, and of different vegetation types such as *low vegetation* and *tree*. The differentiation of impervious cover into *roof* and *pavement* will be hampered by spectral similarities. However, caution is advised when using the reported accuracies as representative quality measures for future EnMAP-based urban data products. They rather provide an estimate for the accuracy at which land cover could be mapped in cities of similar spatial configuration and spectral composition. For urban areas with high-rise buildings or with a higher amount of natural and dust coated materials, estimates on the mapping quality in this Berlin study may be too optimistic (van der Linden and Hostert 2009). In this work, EnMAP's full spectral characteristics, more complex spectral mixtures with more than three materials, and further strategies to improve mappings were not considered during processing. Mapping accuracies therefore represent rather conservative estimates within the limitations mentioned.

2 Main conclusions

Rapid global urbanization constitutes a complex environmental challenge that creates a special situation for remote sensing. Data products that capture urban areas not only by extent, but by their physical surface properties worldwide, could significantly advance our understanding of urban environmental processes (Small 2009). To generate such products, both globally available remote sensing data with a high content of useful information and operational processing schemes that optimally utilize latent information are required. This work addressed these requirements in two aspects: First, the combination of SVR with synthetically mixed training data was introduced as a universal mapping approach that allows the reliable quantification of urban land cover from imaging spectrometer data at multiple spatial scales. Second, the value of hyperspectral imagery from the upcoming satellite mission EnMAP for urban land cover assessments was demonstrated.

Based on the methodological developments of the presented research, the following main conclusions can be drawn:

- *Machine learning-based SVR is well suited for quantifying urban land cover by means of imaging spectrometer data at multiple spatial scales.* By using the concept of kernels and optimal approximating hyperplanes, SVR is capable of coping with urban surface types with high within-class variability and inter-class similarity. At the same time, SVR relativizes the mixed pixel problem within both higher spatial resolution airborne and coarser spatial resolution spaceborne imagery. In the urban context, SVR proved to be superior to other machine learning regressors (e.g., artificial neural networks, random forest) and to MESMA.
- *Synthetically mixed training data generated from a spectral library enable the use of regression approaches for sub-pixel mapping*. When properly generated, sets of pure library spectra, synthetically mixed spectra, and related mixing fractions provide valuable training information for quantifying land cover by means of empirical regression. This approach is straightforward and flexible, and prevents the dependency on a high resolution reference map commonly used for land cover mapping with regression methods.
- *The combination of SVR with synthetically mixed training holds the potential as universal, quantitative mapping approach that will become particularly useful once hyperspectral satellite imagery from urban areas is available.* By using a spectral library as the sole input, SVR models can be trained and flexibly applied to spaceborne

imaging spectrometer data for quantifying urban land cover. The generation of global regression models that are trained on universal spectral databases and generally transferable across space and time may constitute a major step toward the realization of an operational processing scheme for deriving global urban land cover products.

Based on experiments on simulated EnMAP data presented in this research, the following main conclusions can be drawn:

- *The high spectral information content of data acquired by the upcoming hyperspectral satellite mission EnMAP will introduce a new quality into urban remote sensing*. The detection of narrow-band absorption features in visible, near-infrared, and shortwave-infrared wavelengths provides more material-related information and will thus improve the discrimination of urban surface types. This will enhance urban land cover assessments, particularly when compared to those obtained from established multispectral sensors with similar spatial resolution such as Landsat. Such expectancy encourages ongoing developments in operational spaceborne imaging spectrometers for urban applications.
- *The coarser spatial resolution of data acquired by the forthcoming spaceborne imaging spectrometer EnMAP will complicate studies in heterogeneous urban areas*. With a spatial resolution of 30 m, which exceed the 10 to 20 m characteristic scale of individual objects in the urban mosaic (Small 2003), large numbers of urban pixels within EnMAP data will be spectrally mixed. Compared to images from established multispectral systems, mixed pixels within spaceborne imaging spectrometer data contain valuable hyperspectral information which enhances the decomposition of fractions of pure material components. This creates new opportunities for quantitative land cover mapping in environments characterized by a large number of mixed pixels.
- *EnMAP data holds great potential for describing the heterogeneity of urban land cover and may considerably contribute to the generation of new urban data products.* With the launch of the EnMAP mission, urban land cover assessments by means of high quality hyperspectral imagery can move from individual case studies toward global applications. Provided that the spectral information is best utilized, existing maps of urban extent worldwide may be refined by physical surface properties such as the VIS components or beyond. Once a continuous record of EnMAP data is available, EnMAP will allow for monitoring urban land cover over time. This may contribute to regular updates of data products and to the assessment of urban growth patterns.

3 Applications

Data on the amount and spatial distribution of different impervious, pervious, and vegetation cover types are used in a variety of applications that analyze urban environments or that are involved in urban planning. Urban climatology and urban hydrology are two disciplines which utilize such data sets to assess status and changes of urban environmental conditions and functioning. In this context, the role of impervious surfaces as major contributors to environmental impacts and the importance of urban vegetation as a versatile counterpart for mitigation are frequently mentioned. The spatial configuration of built-up components as well as the high amount of impervious surface materials in cities change radiative, thermal, moisture, and aerodynamic characteristics, which ultimately leads to the formation of urban heat islands (Oke 1982; Grimmond 2007). Increased soil sealing lowers infiltration rates and rises surface runoff, which ultimately leads to increased vulnerability to flooding, amplified pollution transport, or reduced ground water recharge (Arnold and Gibbons 1996). Such alterations in climate or hydrosystems are caused by human building activities and can, in turn, influence the quality of life and human health (Harlan et al. 2006; Tan et al. 2010). In contrast, urban vegetation exerts a strong effect on the wind, temperature, moisture, and precipitation regime, and thereby reduces the heat island effect or mitigates air pollution (Avissar 1996). Urban vegetation plays an important role for retarding surface runoff after rainfall events (Sanders 1986). Besides aesthetic and recreational importance, urban dwellers hence benefit from the various ecosystem services urban vegetation provides (Bolund and Hunhammar 1999). Mapping impervious and vegetation from EnMAP data (see Chapter IV) could therefore significantly contribute to research in urban climatology or hydrology. Furthermore, maps could provide urban planners with valuable information necessary for balancing environmental impacts. At what level of detail modelers and decision makers require such information depends on the particular application. Descriptions of impervious or vegetation cover can be used as broad environmental indicators for urban heat islands studies or for hydrological modeling (Weng et al. 2004; Yuan and Bauer 2007; Chormanski et al. 2008). In contrast, descriptions of different impervious and vegetation cover types are important to assess more detailed environmental processes. For example, there are significant differences between different rooftops and paved-areas as well as grass and tree cover types in terms of their thermal response or runoff and infiltration rates (Quattrochi and Ridd 1994; Pauleit and Duhme 2000). The possibility to produce and extend VIS maps underscores the value of EnMAP data for providing such information.

While remote sensing data is commonly used to delineate land cover, the functional characteristics of urban areas or urban land use are often of primary interest for urban planners and modelers (Van de Voorde et al. 2011). Urban structure types (also known as urban morphology types) partition urban areas into distinct spatial units of similar physical and functional characteristics and are a commonly used spatial units in urban planning. The physical components and the various functions they accommodate largely describe the environmental quality of the urban system (Pauleit and Duhme 2000; Gill et al. 2008). The exploration of the use of remote sensing data for operationally mapping and characterizing urban structure types is one ongoing research area in urban remote sensing (Bochow et al. 2010; Van de Voorde et al. 2011; Heiden et al. 2012). The basic assumption of such a mapping relies on the possibility to link characteristic combination of information derived from remote sensing imagery (e.g., land cover, size and arrangement of buildings, etc.) to specific urban structure types (e.g., high- to low-density residential areas, industrial zone, parks, etc.). Fraction maps derived from EnMAP data preserve the heterogeneity of urban land cover and can be related to different urban structure types at the urban block level (Figure V-1). Thus, EnMAP-based data products may be well utilized to link urban land cover information with urban functionality and land use.

Urban information systems are platforms for collecting and exchanging digital information on urban areas that is required for various applications, such as in the mentioned disciplines of urban climatology, hydrology, or urban planning. The value of remote sensing for operationally feeding such systems has been mentioned (Maktav et al. 2005; Small 2006), particularly as an alternative to cost- and time-intensive field surveys or manual mappings from aerial imagery (Jensen and Cowen 1999; Small 2006). The Urban and Environmental Information System of Berlin is one such inventory providing digital data on soil, water, air, climate, land use, traffic, noise, and energy (SenStadt 2014b). For example, the degree imperviousness was derived using a combination of cadastral information and multispectral SPOT-5 data at 10 m spatial resolution, as well as a complex hybrid classification approach (Coenradie and Haag 2012). Results of this work demonstrate the value of future EnMAP data for providing similar descriptions (Figure V-2) and beyond. The high spectral information content of EnMAP enables the use of simple yet powerful analysis concepts, such as the presented SVR approach, on a purely spectral basis. Thus, a considerable contribution to the derivation of reliable and regularly updated urban surface inventories can be made, which might be of even greater value for urban areas where such information systems are not available.

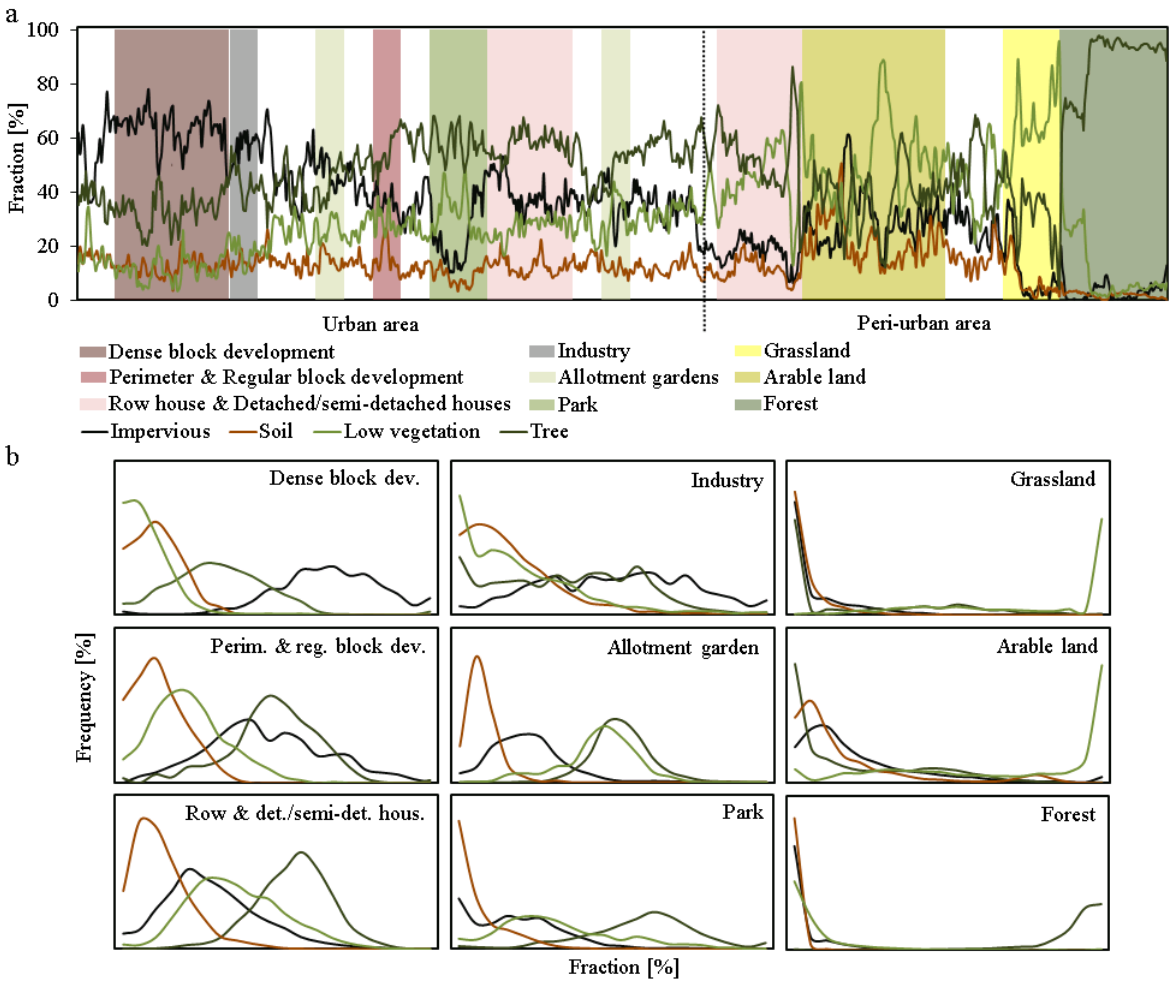


Figure V-1: Relationship between land cover fractions derived from simulated EnMAP data and urban structure types and peri-urban land uses along Berlin's urban-rural gradient. Land cover fractions are presented as (a) spatially continuous values along an N-S transect, and as (b) density graphs. Such density information of urban land cover can be used for automated classification of urban structure types.

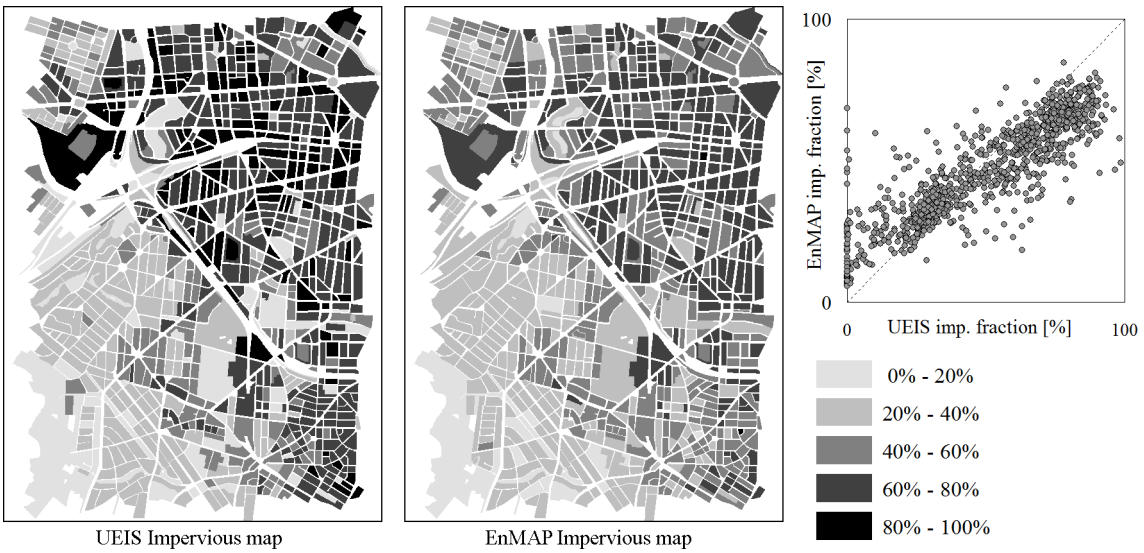


Figure V-2: Comparison between impervious maps from the Berlin Urban and Environmental Information System (UEIS) and from simulated EnMAP data. The high agreement between maps (R^2 = 0.75, MAE = 11.4%) reveal EnMAP's potential for deriving such inventory information on a purely spectral basis. Urban blocks less then 3600 m² (i.e., 4 EnMAP pixels) and urban blocks of traffic infrastructure are excluded.

4 Outlook

In this work, a universal quantitative mapping approach based on SVR and synthetically mixed training data was introduced and the value of data from the upcoming hyperspectral satellite mission EnMAP for urban mapping was demonstrated. Follow-up studies are required to further consolidate and extend these findings. Given the global dimension of urbanization and the need for global urban land cover products, future research should undoubtedly focus on universality and applicability in the broadest possible sense.

With regard to global mapping of urban areas, it is worthwhile to apply the same methodological approach to spaceborne imaging spectrometer data with larger spatial dimensions and to data from different urban areas. This work was carried out along Berlin's urban-rural gradient, a spatially and spectrally heterogeneous study site representative of many cities in general. The investigation of the possibility to map the entire urban area or to map cities of similar spatial configuration and spectral composition would compliment findings and may provide important insights in how to generalize the methodology in order to move beyond isolated case studies. To assess global applicability, validation through mapping of urban areas with different spatial configurations and spectral compositions would be required. Human settlements with high-rise buildings, where entire street canyons will be occluded, or cities in semi-arid to arid ecosystems, where a higher amount of natural and dust coated materials increase spectral similarities (van der Linden and Hostert 2009), may represent challenging and informative study regions for follow-up studies.

From the technical perspective, two major directions could be pursued. First, the proposed SVR approach may be enhanced with respect to the accuracy and reliability of mapping results. Strategies to minimize impacts of spectral similarities and shaded areas typical for urban areas are required. In addition, strategies to cope with large volumes of training data, which will originate from larger input spectral libraries or from increasing the mixing complexity of synthetic mixtures, are needed. To cope with these issues, data transformations prior to modeling (Wu 2004; Feilhauer et al. 2010) or feature selection techniques (Heiden et al. 2007; Rabe et al. 2010b) seem promising techniques. Second, the possibility to train universal SVR models that are generally applicable to arbitrary urban areas may be explored. The concept of using a spectral library as the sole input parameter for synthetically mixed training data generation enables the implementation of such an idea. The spectral library used in this work only comprised material spectra that were relevant for the study region, and SVR models are therefore probably only valid for cities

of similar spectral composition. Universal spectral databases representative of many urban areas and the generation of synthetically mixed training data thereof may be used to make SVR models transferable. This would constitute a major step toward the realization of an operational processing scheme required for deriving urban land cover products globally.

With regard to follow-up applications, analyses may focus on urban mapping assessments that meet the semantic views and requirements of different disciplines. The mapping of urban structure types (compare Section V-3) instead of only land cover may rather fulfill the needs of urban planners. The mapping of surface types in consideration of their thermal properties may help to better describe and understand urban temperature patterns (Gluch et al. 2006). The mapping of environmental indicators that evaluate environmental functions and services may be suited to assess the ecological value of urban areas (Lakes and Kim 2012). Multitemporal EnMAP data may be well utilized to improve descriptions of different urban vegetation types, especially with regard to their role in providing important ecosystems services (Bolund and Hunhammar 1999). In this context, phenological patterns may be used to improve detailed vegetation mapping (Tigges et al. 2013) or as indicators for impacts of urbanization on ecosystem functioning (Zhang et al. 2004).

Once forthcoming hyperspectral satellite missions will be launched, imaging spectrometry holds the potential to become a globally available EO tool contributing to research in Earth System Science. Imaging spectrometer data may considerably help to assess detailed patterns and processes within different ecosystems, and thus complement current and long-term research carried out on multispectral satellite imagery from the Landsat series (Wulder et al. 2012; Roy et al. 2014) or research that will be carried out on forthcoming missions such as Sentinel-2 (Drusch et al. 2012). Given the magnitude of global urbanization and the environmental impacts associated with this process, urban development has to be considered as one of the key human-induced processes driving environmental change from local to global scales (Grimm et al. 2008). Data products derived from remote sensing can be the common link between research, policy, and practice for coping with pressing urban and global change issues (Seto 2009). While global urban mapping assessments have so far delineated urban extent (Potere et al. 2009; Schneider et al. 2010; Esch et al. 2012), future remote sensing research must focus on quantifying urban surface properties within these limits. This work demonstrates the potential of spaceborne imaging spectrometry and machine learning to provide such descriptions.

References

Alberti, M. (2005). The effects of urban patterns on ecosystem function. *International Regional Science Review, 28*, 168-192.

Alonzo, M., Bookhagen, B., & Roberts, D.A. (2014). Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sensing of Environment, 148*, 70-83.

Anderson, P., Okereke, C., Rudd, A., & Parnell, S. (2013). Regional assessment of Africa. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities - A Global Assessment* (pp. 453-459). Dordrecht, Heidelberg, New York, London: Springer.

Antrop, M. (2004). Landscape change and the urbanization process in Europe. *Landscape and Urban Planning, 67*, 9-26.

Arnold, C.L., & Gibbons, C.J. (1996). Impervious surface coverage - The emergence of a key environmental indicator. *Journal of the American Planning Association, 62*, 243-258.

Asner, G.P., & Lobell, D.B. (2000). A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sensing of Environment, 74*, 99-112.

Avissar, R. (1996). Potential effects of vegetation on the urban thermal environment. *Atmospheric Environment, 30*, 437-448.

Bacour, C., Baret, F., Beal, D., Weiss, M., & Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAIxC(ab), from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing of Environment, 105*, 313-325.

Bagan, H., & Yamagata, Y. (2012). Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years. *Remote Sensing of Environment, 127*, 210-222.

Bauer, M.E., Loffelholz, B.C., & Wilson, B. (2008). Estimating and Mapping Impervious Surface Area by Regression Analysis of Landsat Imagery. In Weng, Q. (Ed.) *Remote Sensing of Impervious Surfaces* (pp. 3-19).

Besussi, E., Chin, N., Batty, M., & Longley, P.A. (2010). The structure and form of urban settlements. In Rashed, T. & Jürgens, C. (Eds.), *Remote Sensing of Urban and Suburban Areas* (pp. 13-31). Springer.

Blair, R.B. (1996). Land use and avian species diversity along an urban gradient. *Ecological Applications, 6*, 506-519.

Bochow, M., Taubenbo, x, ck, H., Segl, K., & Kaufmann, H. (2010). An automated and adaptable approach for characterizing and partitioning cities into urban structure types. *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 25-30 July 2010, (pp. 1796-1799).

Bolund, P., & Hunhammar, S. (1999). Ecosystem services in urban areas. *Ecological Economics, 29*, 293-301.

Borel, C.C., & Gerstl, S.A.W. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment, 47*, 403-416.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-32.

Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.

Brereton, R.G., & Lloyd, G.R. (2010). Support vector machines for classification and regression. *Analyst, 135*, 230-267.

Burges, C.J.C. (1998). A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*, 121-167.

Cadenasso, M.L., Pickett, S.T.A., & Schwarz, K. (2007). Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification. *Frontiers in Ecology and the Environment, 5*, 80-88.

Caicedo, J.P.R., Verrelst, J., Munoz-Mari, J., Moreno, J., & Camps-Valls, G. (2014). Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, PP*, 1-1.

Camps-Valls, G., & Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 43*, 1351-1362.

Camps-Valls, G., & Bruzzone, L. (2009). *Kernel Methods for Remote Sensing Data Analysis*.Wiley & Sons.

Camps-Valls, G., Bruzzone, L., Rojo-Alvarez, J.L., & Melgani, F. (2006). Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geoscience and Remote Sensing Letters, 3*, 339-343.

Cavalli, R.M., Fusilli, L., Pascucci, S., Pignatti, S., & Santini, F. (2008). Hyperspectral sensor data capability for retrieving complex urban land cover in comparison with multispectral data: Venice city case study (Italy). *Sensors, 8*, 3299-3320.

Cernicharo, J., Verger, A., & Camacho, F. (2013). Empirical and Physical Estimation of Canopy Water Content from CHRIS/PROBA Data. *Remote Sensing, 5*, 5265-5284.

Chan, J.C.-W., & Paelinckx, D. (2008). Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment, 112*, 2999-3011.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2*.

Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems, 78*, 103-112.

Chormanski, J., Van de Voorde, T., De Roeck, T., Batelaan, O., & Canters, F. (2008). Improving distributed runoff prediction in urbanized catchments with remote sensing based estimates of impervious surface cover. *Sensors, 8*, 910-932.

Chudnovsky, A., Ben-Dor, E., & Saaroni, H. (2004). Diurnal thermal behavior of selected urban objects using remote sensing measurements. *Energy and Buildings, 36*, 1063-1074.

Chuvieco, E. (2008). *Earth Observation of Global Change - The Role of Satellite Remote Sensing in Monitoring the Global Environment*.Springer.

Clark, M.L., & Roberts, D.A. (2012). Species-Level Differences in Hyperspectral Metrics among Tropical Rainforest Trees as Determined by a Tree-Based Classifier. *Remote Sensing, 4*, 1820-1855.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., & Shields, T. (1998). The HyMap™ airborne hyperspectral sensor: The system, calibration and performance. *1st Earsel Workshop on Imaging Spectroscopy*, 37-42.

Coenradie, B., & Haag, L. (2012). *Versiegelungskartierung Berlin - Anwendung und Weiterentwicklung des hybriden Auswertungsverfahrens für das Jahr 2011 sowie Kartierung von Veränderungen. Abschlussbericht. Senatsverwaltung für Stadtentwicklung*. Senatsverwaltung für Stadtentwicklung Berlin.

Colgan, M.S., Baldeck, C.A., Feret, J.-B., & Asner, G.P. (2012). Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data. *Remote Sensing, 4*, 3462-3480.

DeFries, R.S., Foley, J.A., & Asner, G.P. (2004). Land-use choices: balancing human needs and ecosystem function. *Frontiers in Ecology and the Environment, 2*, 249-257.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment, 120*, 25-36.

Durbha, S.S., King, R.L., & Younan, N.H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment, 107*, 348-361.

Esch, T., Himmler, V., Schorcht, G., Thiel, M., Wehrmann, T., Bachofer, F., Conrad, C., Schmidt, M., & Dech, S. (2009). Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. *Remote Sensing of Environment, 113*, 1678-1690.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenboeck, H., Mueller, A., & Dech, S. (2013). Urban Footprint Processor-Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geoscience and Remote Sensing Letters, 10*, 1617-1621.

Esch, T., Taubenboeck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., Schmidt, M., Mueller, A., & Dech, S. (2012). TanDEM-X mission-new perspectives for the

inventory and monitoring of global settlement patterns. *Journal of Applied Remote Sensing, 6*.

Feilhauer, H., Asner, G.P., Martin, R.E., & Schmidtlein, S. (2010). Brightness-normalized Partial Least Squares Regression for hyperspectral data. *Journal of Quantitative Spectroscopy and Radiative Transfer, 111*, 1947-1957.

Fenger, J. (1999). Urban air quality. *Atmospheric Environment, 33*, 4877-4900.

Folke, C., Jansson, A., Larsson, J., & Costanza, R. (1997). Ecosystem appropriation by cities. *Ambio, 26*, 167-172.

Fontana, S., Sattler, T., Bontadina, F., & Moretti, M. (2011). How to manage the urban green to improve bird diversity and community structure. *Landscape and Urban Planning, 101*, 278-285.

Foody, G.M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*, 1335-1343.

Foody, G.M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment, 103*, 179-189.

Fragkias, M., Güneralp, B., Seto, K.C., & Goodness, J. (2013). A synthesis of global urbanization projections. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities - A Global Assessment* (pp. 409-435). Dordrecht, Heidelberg, New York, London: Springer.

Franke, J., Roberts, D.A., Halligan, K., & Menz, G. (2009). Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment, 113*, 1712-1723.

Friedl, M.A., & Brodley, C.E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment, 61*, 399-409.

Galeazzi, C., Sacchetti, A., Cisbani, A., & Babini, G. (2008). The PRISMA Program. *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, 7-11 July 2008, 4 (pp. IV - 105-IV - 108).

Gamba, P., & Dell'Acqua, F. (2006). Spectral resolution in the context of very high resolution urban remote sensing. In Quattrochi, D.W., Q. (Ed.) *Urban Remote Sensing* (pp. 377-391). New York: CRC Press Inc.

Gill, S.E., Handley, J.F., Ennos, A.R., Pauleit, S., Theuray, N., & Lindley, S.J. (2008). Characterising the urban environment of UK cities and towns: A template for landscape planning. *Landscape and Urban Planning, 87*, 210-222.

Gluch, R., Quattrochi, D.A., & Luvall, J.C. (2006). A multi-scale approach to urban thermal analysis. *Remote Sensing of Environment, 104*, 123-132.

Goetz, A.F.H. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment, 113*, S5-S16.

Green, R.O., Eastwood, M.L., Sarture, C.M., Chrien, T.G., Aronsson, M., Chippendale, B.J., Faust, J.A., Pavri, B.E., Chovit, C.J., Solis, M., Olah, M.R., & Williams, O. (1998). Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment, 65*, 227-248.

Griffiths, P., Hostert, P., Gruebner, O., & van der Linden, S. (2010). Mapping megacity growth with multi-sensor data. *Remote Sensing of Environment, 114*, 426-439.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J.G., Bai, X.M., & Briggs, J.M. (2008). Global change and the ecology of cities. *Science, 319*, 756-760.

Grimmond, S. (2007). Urbanization and global environmental change: local effects of urban warming. *Geographical Journal, 173*, 83-88.

Guanter, L., Segl, K., & Kaufmann, H. (2009). Simulation of Optical Remote-Sensing Scenes With Application to the EnMAP Hyperspectral Mission. *Geoscience and Remote Sensing, IEEE Transactions on, 47*, 2340-2351.

Hahs, A.K., & McDonnell, M.J. (2006). Selecting independent measures to quantify Melbourne's urban-rural gradient. *Landscape and Urban Planning, 78*, 435-448.

Harlan, S.L., Brazel, A.J., Prashad, L., Stefanov, W.L., & Larsen, L. (2006). Neighborhood microclimates and vulnerability to heat stress. *Social Science & Medicine, 63*, 2847-2863.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*.Springer

Haykin, S. (1999). *Neural Networks - A Comprehensive Foundation*.Prentice Hall.

Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., & Mueller, A. (2012). Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and Urban Planning, 105*, 361-375.

Heiden, U., Segl, K., Roessner, S., & Kaufmann, H. (2007). Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sensing of Environment, 111*, 537-552.

Held, M., Hostert, P., Jakimow, B., van der Linden, S., Okujeni, A., Rabe, A., Suess, S., & Wirth, F. (2014). *EnMAP-Box Documentation Center, Version 2.0.2*. Germany: Humboldt-Universität zu Berlin.

Heldens, W., Heiden, U., Esch, T., Stein, E., & Mueller, A. (2011). Can the future EnMAP mission contribute to urban applications? A literature survey. *Remote Sensing, 3*, 1817-1846.

Herold, M. (2009). Some Recommendations for Global Efforts in Urban Monitoring and Assessments from Remote Sensing. In Gamba, P. & Herold, M. (Eds.), *Global Mapping of Human Settlement* (pp. 11-23). CRC Press.

Herold, M., Gardner, M.E., & Roberts, D.A. (2003). Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 1907-1919.

Herold, M., Roberts, D.A., Gardner, M.E., & Dennison, P.E. (2004). Spectrometry for urban area remote sensing - Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment, 91*, 304-319.

Herold, M., Schiefer, S., Hostert, P., & Roberts, D.A. (2006). Applying imaging spectrometry in urban areas. In Quattrochi, D. & Weng, Q. (Eds.), *Urban Remote Sensing* (pp. 137-161). New York: CRC Press Inc.

Huang, C., Davis, L.S., & Townshend, J.R.G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing, 23*, 725-749.

Im, J., Jensen, J.R., Jensen, R.R., Gladden, J., Waugh, J., & Serrato, M. (2012a). Vegetation Cover Analysis of Hazardous Waste Sites in Utah and Arizona Using Hyperspectral Remote Sensing. *Remote Sensing, 4*, 327-353.

Im, J., Lu, Z., Rhee, J., & Quackenbush, L.J. (2012b). Impervious surface quantification using a synthesis of artificial immune networks and decision/regression trees from multi-sensor data. *Remote Sensing of Environment, 117*, 102-113.

Imhoff, M.L., Zhang, P., Wolfe, R.E., & Bounoua, L. (2010). Remote sensing of the urban heat island effect across biomes in the continental USA. *Remote Sensing of Environment, 114*, 504-513.

Jensen, J.R., & Cowen, D.C. (1999). Remote sensing of urban suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing, 65*, 611-622.

Kareiva, P., Watts, S., McDonald, R., & Boucher, T. (2007). Domesticated nature: Shaping landscapes and ecosystems for human welfare. *Science, 316*, 1866-1869.

Kaufmann, H., Segl, K., Guanter, L., Hofer, S., Foerster, K.P., Stuffler, T., Mueller, A., Richter, R., Bach, H., Hostert, P., & Chlebek, C. (2008). Environmental Mapping and Analysis Program (EnMAP) - Recent Advances and Status. *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, 7-11 July 2008, 4 (pp. IV - 109-IV - 112).

Kennedy, C., Steinberger, J., Gasson, B., Hansen, Y., Hillman, T., Havranek, M., Pataki, D., Phdungsilp, A., Ramaswami, A., & Villalba Mendez, G. (2009). Greenhouse Gas Emissions from Global Cities. *Environmental Science & Technology, 43*, 7297-7302.

Kong, F., & Nakagoshi, N. (2006). Spatial-temporal gradient analysis of urban green spaces in Jinan, China. *Landscape and Urban Planning, 78*, 147-164.

Kronenberg, J., Tezer, A., Haase, D., & Colding, J. (2013). Regional assessment of Europe. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization, Biodiversity and Ecosystem Services: Challenges and*

Opportunities - A Global Assessment (pp. 275-278). Dordrecht, Heidelberg, New York, London: Springer.

Kruse, F.A., Boardman, J.W., & Huntington, J.F. (2003). Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 1388-1400.

Lakes, T., & Kim, H.-O. (2012). The urban environmental indicator "Biotope Area Ratio"-An enhanced approach to assess and manage the urban ecosystem services using high resolution remote-sensing. *Ecological Indicators, 13*, 93-103.

Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skånes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., & Xu, J. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change, 11*, 261-269.

Laurent, V.C.E., Schaepman, M.E., Verhoef, W., Weyermann, J., & Chávez, R.O. (2014). Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance image. *Remote Sensing of Environment, 140*, 318-329.

Liang, S. (2008). *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*.Springer.

Lu, D., Li, G., Moran, E., Batistella, M., & Freitas, C.C. (2011). Mapping impervious surfaces with the integrated use of Landsat Thematic Mapper and radar data: A case study in an urban–rural landscape in the Brazilian Amazon. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*, 798-808.

Maktav, D., Erbek, F.S., & Jurgens, C. (2005). Remote sensing of urban areas. *International Journal of Remote Sensing, 26*, 655-659.

Masser, I. (2001). Managing our urban future: the role of remote sensing and geographic information systems. *Habitat International, 25*, 503-512.

Matsunaga, T., Iwasaki, A., Tsuchida, S., Tanii, J., Kashimura, O., Nakamura, R., Yamamoto, H., Tachikawa, T., & Rokugawa, S. (2013). Current status of Hyperspectral Imager Suite (HISUI). *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, 21-26 July 2013, (pp. 3510-3513).

McDonnell, M.J., & Pickett, S.T.A. (1990). Ecosystem Structure and Function along Urban-Rural Gradients: An Unexploited Opportunity for Ecology. *Ecology, 71*, 1232-1237.

McKinney, M.L. (2002). Urbanization, biodiversity, and conservation. *Bioscience, 52*, 883-890.

McPhearson, T., Auch, R., & Alberti, M. (2013). Regional assessment of North America - Urbanization trends, biodiversity patterns, and ecosystem services. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization,*

Biodiversity and Ecosystem Services: Challenges and Opportunities - A Global Assessment (pp. 279-286). Dordrecht, Heidelberg, New York, London: Springer.

MEA [Millenium Ecosystem Assessment] (2005). *Ecosystems and Human Well-being: Synthesis*. Washington, DC: World Resources Institute.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*, 1778-1790.

Middel, A., Haeb, K., Brazel, A.J., Martin, C.A., & Guhathakurta, S. (2014). Impact of urban form and design on mid-afternoon microclimate in Phoenix Local Climate Zones. *Landscape and Urban Planning, 122*, 16-28.

Miller, R.B., & Small, C. (2003). Cities from space: Potential applications of remote sensing in urban environmental research and policy. *Environmental Science & Policy, 6*, 129-137.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*, 247-259.

Myeong, S., Nowak, D.J., & Duggin, M.J. (2006). A temporal analysis of urban forest carbon storage using remote sensing. *Remote Sensing of Environment, 101*, 277-282.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*, 1145-1161.

NRC [National Research Council] (1999). *Our Common Journey: A Transition Toward Sustainability*.The National Academies Press.

NRC [National Research Council] (2007). *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*.The National Academies Press.

Oke, T.R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society, 108*, 1-24.

Okin, G.S., Roberts, D.A., Murray, B., & Okin, W.J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment, 77*, 212-225.

Okujeni, A., van der Linden, S., Jakimow, B., Rabe, A., Verrelst, J., & Hostert, P. (2014). A Comparison of Advanced Regression Algorithms for Quantifying Urban Land Cover. *Remote Sensing, 6*, 6324-6346.

Okujeni, A., van der Linden, S., Tits, L., Somers, B., & Hostert, P. (2013). Support vector regression and synthetically mixed training data for quantifying urban land cover. *Remote Sensing of Environment, 137*, 184-197.

Pal, M., & Mather, P.M. (2006). Some issues in the classification of DAIS hyperspectral data. *International Journal of Remote Sensing, 27*, 2895-2916.

Pauchard, A., Barbosa, O., Maira, J., Rojas, C., Villagra, P., Faggi, A., Márquez, F., Aponte, G., & MacGregor-Fors, I. (2013). Regional assessment of Latin America: Rapid urban development and social-economic inequity threaten biodiversity hotspots. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities - A Global Assessment* (pp. 589-608). Dordrecht, Heidelberg, New York, London: Springer.

Pauleit, S., & Duhme, F. (2000). Assessing the environmental performance of land cover types for urban planning. *Landscape and Urban Planning, 52*, 1-20.

Pelling, M. (2003). *The Vulnerability of Cities: Natural Disasters and Social Resilience*. London: Earthscan Publications Ltd.

Pesaresi, M., & Ehrlich, D. (2009). A Methodology to Quantify Built-Up Structures from Optical VHR Imagery. In Gamba, P. & Herold, M. (Eds.), *Global Mapping of Human Settlement* (pp. 27-58). CRC Press.

Pickett, S.T.A., Cadenasso, M.L., Grove, J.M., Boone, C.G., Groffman, P.M., Irwin, E., Kaushal, S.S., Marshall, V., McGrath, B.P., Nilon, C.H., Pouyat, R.V., Szlavecz, K., Troy, A., & Warren, P. (2011). Urban ecological systems: Scientific foundations and a decade of progress. *Journal of Environmental Management, 92*, 331-362.

Potere, D., Schneider, A., Angel, S., & Civco, D.L. (2009). Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *International Journal of Remote Sensing, 30*, 6531-6558.

Powell, R.L., Roberts, D.A., Dennison, P.E., & Hess, L.L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment, 106*, 253-267.

Pu, R., Gong, P., Michishita, R., & Sasagawa, T. (2008). Spectral mixture analysis for mapping abundance of urban surface components from the Terra/ASTER data. *Remote Sensing of Environment, 112*, 939-954.

Pu, R., & Landry, S. (2012). A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment, 124*, 516-533.

Quattrochi, D.A., & Ridd, M.K. (1994). Measurement and analysis of thermal energy responses from discrete urban surfaces using remote sensing data. *International Journal of Remote Sensing, 15*, 1991-2022.

Quigley, J.M. (1998). Urban Diversity and Economic Growth. *Journal of Economic Perspectives, 12*, 127-138.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning, 1*, 81-106. In English.

Rabe, A., Jakimow, B., van der Linden, S., & Hostert, P. (2012). EnMAP Box, Version 1.4 [online]. Available from: http://indus.caf.dlr.de/forum/ [accessed April 2013].

Rabe, A., Jakimow, B., van der Linden, S., Okujeni, A., Suess, S., Leitao, P.J., & Hostert, P. (2013). Simplifying Support Vector Regression Parameterisation by Heuristic Search for Optimal e-Loss. *8th EARSeL Workshop of Special Interest Group in Imaging Spectroscopy. Nantes, France*.

Rabe, A., van der Linden, S., & Hostert, P. (2010a). imageSVM, Version 2.1. [online]. Available from: http://www.imagesvm.net/ [accessed March 2014].

Rabe, A., van der Linden, S., & Hostert, P. (2010b). Simplifying Support Vector Machines for classification of hyperspectral imagery and selection of relevant features. *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on*, 14-16 June 2010, (pp. 1-4).

Rabe, A., van der Linden, S., & Hostert, P. (2014a). EnMAP Box, Version 2.0 [online]. Available from: http://www.enmap.org/ [accessed July 2014].

Rabe, A., van der Linden, S., & Hostert, P. (2014b). imageSVM, Version 3.0 [online]. Available from: http://www.imagesvm.net/ [accessed July 2014].

Rashed, T., Weeks, J.R., Roberts, D., Rogan, J., & Powell, R. (2003). Measuring the physical composition of urban morphology using multiple endmember spectral mixture models. *Photogrammetric Engineering and Remote Sensing, 69*, 1011-1020.

Rees, W.E. (1992). Ecological footprints and appropriated carrying capacity: What urban economics leaves out. *Environment and Urbanization, 4*, 121-130.

Richter, R., & Schläpfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. *International Journal of Remote Sensing, 23*, 2631-2649.

Ridd, M.K. (1995). Exploring a V-I-S (Vegetation-Impervious surface-Soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities. *International Journal of Remote Sensing, 16*, 2165-2185.

Roberts, D.A., Dennison, P.E., Gardner, M.E., Hetzel, Y., Ustin, S.L., & Lee, C.T. (2003). Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible/Infrared Imaging Spectrometer. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 1297-1310.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R.O. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment, 65*, 267-279.

Roberts, D.A., Quattrochi, D.A., Hulley, G.C., Hook, S.J., & Green, R.O. (2012). Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission. *Remote Sensing of Environment, 117*, 83-101.

Roberts, D.A., Smith, M.O., & Adams, J.B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment, 44*, 255-269.

Roessner, S., Segl, K., Heiden, U., & Kaufmann, H. (2001). Automated differentiation of urban surfaces based on airborne hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing, 39*, 1525-1532.

Roy, D.P., Wulder, M.A., Loveland, T.R., C.E, W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment, 145*, 154-172.

Sánchez-Rodríguez, R., Seto, K.C., Simon, D., Solecki, W.D., Kraas, F., & Laumann, G. (2005). *Science Plan: Urbanization and Global Environmental Change*, IHDP Report Series. Bonn: International Human Dimensions Programme on Global Environmental Change (IHDP)

Sanders, R.A. (1986). Urban vegetation impacts on the hydrology of Dayton, Ohio. *Urban Ecology, 9*, 361-376.

Satterthwaite, D. (2007). *The transition to a pre-dominantly urban world and its underpinnings*, Human Settlements Discussion Paper – Urban Change 4. London: International Institute for Environment and Development (IIED).

Schiefer, S., Hostert, P., & Damm, A. (2006). Correcting brightness gradients in hyperspectral data from urban areas. *Remote Sensing of Environment, 101*, 25-37.

Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: Parametric orthorectification. *International Journal of Remote Sensing, 23*, 2609-2630.

Schmidtlein, S., Feilhauer, H., & Bruelheide, H. (2012). Mapping plant strategy types using remote sensing. *Journal of Vegetation Science, 23*, 395-405.

Schneider, A., Friedl, M.A., & Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment, 114*, 1733-1746.

Schölkopf, B., & Smola, A.J. (2002). *Learning with Kernels - Support Vector Machines, Regularization, Optimization, and Beyond.* Cambridge, Massachusetts: MIT Press.

Schulz, M. (2000). Berlin - recent issues in urban development. *Beiträge zur Regionalen Geographie, 52*, 224-236.

Segl, K., Guanter, L., Rogass, C., Kuester, T., Roessner, S., Kaufmann, H., Sang, B., Mogulsky, V., & Hofer, S. (2012). EeteS-The EnMAP End-to-End Simulation Tool. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*, 522-530.

SenStadt [Senatsverwaltung für Stadtentwicklung] (2013a). Berlin Urban and Environmental Information System (UEIS) [online]. Available from: www.stadtentwicklung.berlin.de/umwelt/umweltatlas [accessed April 2013].

SenStadt [Senatsverwaltung für Stadtentwicklung] (2013b). Automatisierte Liegenschaftskarte (ALK-Berlin) [online]. Available from: www.stadtentwicklung.berlin.de/geoinformation/liegenschaftskataster/ [accessed April 2013].

SenStadt [Senatsverwaltung für Stadtentwicklung] (2014a). Automatisierte Liegenschaftskarte (ALK-Berlin) [online]. Available from: www.stadtentwicklung.berlin.de/geoinformation/liegenschaftskataster/ [accessed July 2014].

SenStadt [Senatsverwaltung für Stadtentwicklung] (2014b). Berlin Urban and Environmental Information System (UEIS) [online]. Available from: www.stadtentwicklung.berlin.de/umwelt/umweltatlas [accessed July 2014].

SenStadt [Senatsverwaltung für Stadtentwicklung] (2014c). Urban green space [online]. Available from: http://www.stadtentwicklung.berlin.de/umwelt/stadtgruen/index_en.shtml [accessed June 2014].

Seto, K.C. (2009). Global Urban Issues. In Gamba, P. & Herold, M. (Eds.), *Global Mapping of Human Settlement* (pp. 3-9). CRC Press.

Seto, K.C. (2013). Regional assessment of Asia. In Elmqvist, T., Fragkias, M., Goodness, J., Güneralp, B., Marcotullio, P.J., McDonald, R.I., Parnell, S., Schewenius, M., Sendstad, M., Seto, K.C. & Wilkinson, C. (Eds.), *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities - A Global Assessment* (pp. 53-56). Dordrecht, Heidelberg, New York, London: Springer.

Seto, K.C., Sanchez-Rodriguez, R., & Fragkias, M. (2010). The New Geography of Contemporary Urbanization and the Environment. In Gadgil, A. & Liverman, D.M. (Eds.), *Annual Review of Environment and Resources, Vol 35* (pp. 167-194).

Shackelford, A.K., & Davis, C.H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 2354-2363.

Sjöman, J.D., & Gill, S.E. (2014). Residential runoff – The role of spatial density and surface cover, with a case study in the Höjeå river catchment, southern Sweden. *Urban Forestry & Urban Greening, 13*, 304-314.

Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing, 22*, 1305-1334.

Small, C. (2003). High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment, 88*, 170-186.

Small, C. (2004). The Landsat ETM+ spectral mixing space. *Remote Sensing of Environment, 93*, 1-17.

Small, C. (2006). Comparative analysis of urban reflectance and surface temperature. *Remote Sensing of Environment, 104*, 168-189.

Small, C. (2009). The Color of Cities An Overview of Urban Spectral Diversity. In Herold, P.G.a.M. (Ed.) *Global Mapping of Human Settlement: Experiences, Datasets and Prospects* (pp. 59 - 105). Boca Raton, London, New York: Taylor and Francis Group, LLC.

Small, C., & Lu, J.W.T. (2006). Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment, 100*, 441-456.

Smith, M.L., Martin, M.E., Plourde, L., & Ollinger, S.V. (2003). Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: Comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 1332-1337.

Smola, A.J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing, 14*, 199-222.

Somers, B., Asner, G.P., Tits, L., & Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment, 115*, 1603-1616.

Somers, B., Cools, K., Delalieux, S., Stuckens, J., van der Zande, D., Verstraeten, W.W., & Coppin, P. (2009a). Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sensing of Environment, 113*, 1183-1193.

Somers, B., Delalieux, S., Stuckens, J., Verstraeten, W.W., & Coppin, P. (2009b). A weighted linear spectral mixture analysis approach to address endmember variability in agricultural production systems. *International Journal of Remote Sensing, 30*, 139-147.

Staenz, K., & Held, A. (2012). Summary of current and future terrestrial civilian hyperspectral spaceborne systems. *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, 22-27 July 2012, (pp. 123-126).

Stehman, S.V., & Wickham, J.D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment, 115*, 3044-3055.

Stuffler, T., Förster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H., Penné, B., Mueller, A., & Chlebek, C. (2009). Hyperspectral imaging - An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). *Acta Astronautica, 65*, 1107-1112.

Stuffler, T., Kaufmann, C., Hofer, S., Förster, K.P., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penné, B., Benz, U., & Haydn, R. (2007). The EnMAP hyperspectral imager - An advanced optical payload for future applications in Earth observation programmes. *Acta Astronautica, 61*, 115-120.

Sukopp, H. (1990). *Stadtökologie: Das Beispiel Berlin*. Berlin: Dietrich Reimer Verlag. In German.

Tan, J., Zheng, Y., Tang, X., Guo, C., Li, L., Song, G., Zhen, X., Yuan, D., Kalkstein, A.J., Li, F., & Chen, H. (2010). The urban heat island and its impact on heat waves and human health in Shanghai. *International Journal of Biometeorology, 54*, 75-84.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment, 117*, 162-176.

Thomas, N., Hendrix, C., & Congalton, R.G. (2003). A comparison of urban mapping methods using high-resolution digital imagery. *Photogrammetric Engineering and Remote Sensing, 69*, 963-972.

Tigges, J., Lakes, T., & Hostert, P. (2013). Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment, 136*, 66-75.

Tompkins, S., Mustard, J.F., Pieters, C.M., & Forsyth, D.W. (1997). Optimization of endmembers for spectral mixture analysis. *Remote Sensing of Environment, 59*, 472-489.

Tuia, D., & Camps-Valls, G. (2011). Urban image classification with semisupervised multiscale cluster kernels. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4*, 65-74.

Tuia, D., Verrelst, J., Alonso, L., Perez-Cruz, F., & Camps-Valls, G. (2011). Multioutput support vector regression for remote sensing biophysical parameter estimation. *IEEE Geoscience and Remote Sensing Letters, 8*, 804-808.

UN [United Nations] Habitat (2010). *State of the World's Cities 2010/2011 - Cities for All: Bridging the Urban Divide*

United Nations (2006). *World Urbanization Prospects: The 2005 Revision*. New York:

United Nations (2012a). *World Urbanization Prospects: The 2011 Revision*. New York:

United Nations (2012b). *World Urbanization Prospects: The 2011 Revision - Data on Cities and Urban Agglomerations: 17b. Number of cities classified by size class of urban settlement, 1950-2025*. New York:

Van de Voorde, T., De Roeck, T., & Canters, F. (2009). A comparison of two spectral mixture modelling approaches for impervious surface mapping in urban areas. *International Journal of Remote Sensing, 30*, 4785-4806.

Van de Voorde, T., Jacquet, W., & Canters, F. (2011). Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning, 102*, 143-155.

van der Linden, S., & Hostert, P. (2009). The influence of urban structures on impervious surface maps from airborne hyperspectral data. *Remote Sensing of Environment, 113*, 2298-2305.

van der Linden, S., Janz, A., Waske, B., Eiden, M., & Hostert, P. (2007). Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. *Journal of Applied Remote Sensing, 1*.

van der Linden, S., Rabe, A., Held, A., Wirth, F., Suess, S., Okujeni, A., & Hostert, P. (2014). *imageSVM Regression, Manual for Application: imageSVM version 3.0*. Humboldt-Universität zu Berlin, Germany.

Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer.

Verrelst, J., Munoz, J., Alonso, L., Delegido, J., Pablo Rivera, J., Camps-Valls, G., & Moreno, J. (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment, 118*, 127-139.

Vitousek, P.M., Mooney, H.A., Lubchenco, J., & Melillo, J.M. (1997). Human domination of Earth's ecosystems. *Science, 277*, 494-499.

Walton, J.T. (2008). Subpixel urban land cover estimation: Comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering and Remote Sensing, 74*, 1213-1222.

Waske, B., Fauvel, M., Benediktsson, J.A., & Chanussot, J. (2009). Machine Learning Techniques in Remote Sensing Data Analysis. In Camps-Valls, G. & Bruzzone, L. (Eds.), *Kernel Methods for Remote Sensing Data Analysis*. Chichester, UK: John Wiley & Sons, Ltd.

Waske, B., van der Linden, S., Benediktsson, J.A., Rabe, A., & Hostert, P. (2010). Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 48*, 2880-2889.

Waske, B., van der Linden, S., Oldenburg, C., Jakimow, B., Rabe, A., & Hostert, P. (2012). imageRF – A user-oriented implementation for remote sensing image analysis with Random Forests. *Environmental Modelling & Software, 35*, 192-193.

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 335-344.

Weng, Q. (2011). *Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications*.CRC Press, Taylor & Francis Group.

Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment, 117*, 34-49.

Weng, Q. (2014). What is special about global urban remote sensing? In Weng, Q. (Ed.) *Global urban monitoring and assessment through earth observation* (pp. 1-12). CRC Press, Taylor & Francis Group.

Weng, Q., Hu, X., & Lu, D. (2008). Extracting impervious surfaces from medium spatial resolution multispectral and hyperspectral imagery: a comparison. *International Journal of Remote Sensing, 29*, 3209-3232.

Weng, Q.H., Lu, D.S., & Schubring, J. (2004). Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment, 89*, 467-483.

Westad, F., & Martens, H. (2000). Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. *Journal of near Infrared Spectroscopy, 8*, 117-124.

Wold, S., Sjostrom, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*, 109-130.

Wu, C. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment, 93*, 480-492.

Wu, C.S., & Murray, A.T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment, 84*, 493-505.

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., & Woodcock, C.E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment, 122*, 2-10.

Yang, L.M., Xian, G., Klaver, J.M., & Deal, B. (2003). Urban land-cover change detection through sub-pixel imperviousness mapping using remotely sensed data. *Photogrammetric Engineering and Remote Sensing, 69*, 1003-1010.

Yu, K., Leufen, G., Hunsche, M., Noga, G., Chen, X., & Bareth, G. (2013). Investigation of Leaf Diseases and Estimation of Chlorophyll Concentration in Seven Barley Varieties Using Fluorescence and Hyperspectral Indices. *Remote Sensing, 6*, 64-86.

Yu, X.J., & Ng, C.N. (2007). Spatial and temporal dynamics of urban sprawl along two urban-rural transects: A case study of Guangzhou, China. *Landscape and Urban Planning, 79*, 96-109.

Yuan, F., & Bauer, M.E. (2007). Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. *Remote Sensing of Environment, 106*, 375-386.

Yuan, F., Sawaya, K.E., Loeffelholz, B.C., & Bauer, M.E. (2005). Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment, 98*, 317-328.

Yuan, F., Wu, C., & Bauer, M.E. (2008). Comparison of spectral analysis techniques for impervious surface estimation using Landsat imagery. *Photogrammetric Engineering and Remote Sensing, 74*, 1045-1055.

Zhang, X.Y., Friedl, M.A., Schaaf, C.B., Strahler, A.H., & Schneider, A. (2004). The footprint of urban climates on vegetation phenology. *Geophysical Research Letters, 31*.

Zhou, X., & Wang, Y.-C. (2011). Spatial–temporal dynamics of urban green space in response to rapid urbanization and greening policies. *Landscape and Urban Planning, 100*, 268-277.

Appendix:
synthMix-SVR

Manual for Application: synthMix-SVR (1.0), 14. July 2014

Akpona Okujeni, Marcel Schwieder, Matthias Held, Benjamin Jakimow, Andreas Rabe, Pedro J. Leitão, Sebastian van der Linden and Patrick Hostert

© Humboldt-Universität zu Berlin, Geomatics Lab, 2014.

Copyright for application and manual

© Humboldt-Universität zu Berlin, Geomatics Lab, 2014, www.geographie.hu-berlin.de/labs/geomatics

Please cite this manual as: Okujeni, A., Schwieder, M., Held, M., Jakimow, B., Rabe, A., Leitão, P.J., van der Linden, S., Hostert, P., (2014). synthMix-SVR, Manual for Application: synthMix-SVR version 1.0. Humboldt-Universität zu Berlin, Germany

Please cite this application as: Okujeni, A., Schwieder, M., Held, M., Jakimow, B., Rabe, A., Leitão, P.J., van der Linden, S., Hostert, P., (2014). synthMix-SVR, Version 1.0, software available at www.enmap.org

Disclaimer

The authors of this manual accept no responsibility for errors or omissions in this work and shall not be liable for any damage caused by these errors or omissions..

1 Concept

synthMix-SVR is an IDL-based tool for quantitative analysis of remote sensing data. It implements the concept of combining support vector regression (SVR) with synthetically mixed training data for mapping sub-pixel fractions of land cover (Okujeni et al. 2013). synthMix-SVR is embedded into the EnMAP-Box 2.0 (available from: www.enmap.org) and makes use of imageSVM 3.0 (available from: www.imagesvm.net) for SVR modeling.

The goal of synthMix-SVR is to provide a user-friendly tool for producing land cover fraction maps from remote sensing imagery. synthMix-SVR follows the approach proposed in Okujeni et al. (2013):

(1) A spectral library is used to generate a synthetically mixed training data set for a single land cover category of interest, a so-called target category.

(2) The synthetic data set is used to train a SVR model, which is subsequently used to derive a fraction map of the respective target category.

Building on this concept, synthMix-SVR combines (1) and (2) and additionally allows the user to flexibly map multiple target categories through iterative processing (Figure A-1). Thus, synthMix-SVR is suited for a comprehensive mapping of land cover, delivering a set of multiple target category fraction maps in a multilayer image.

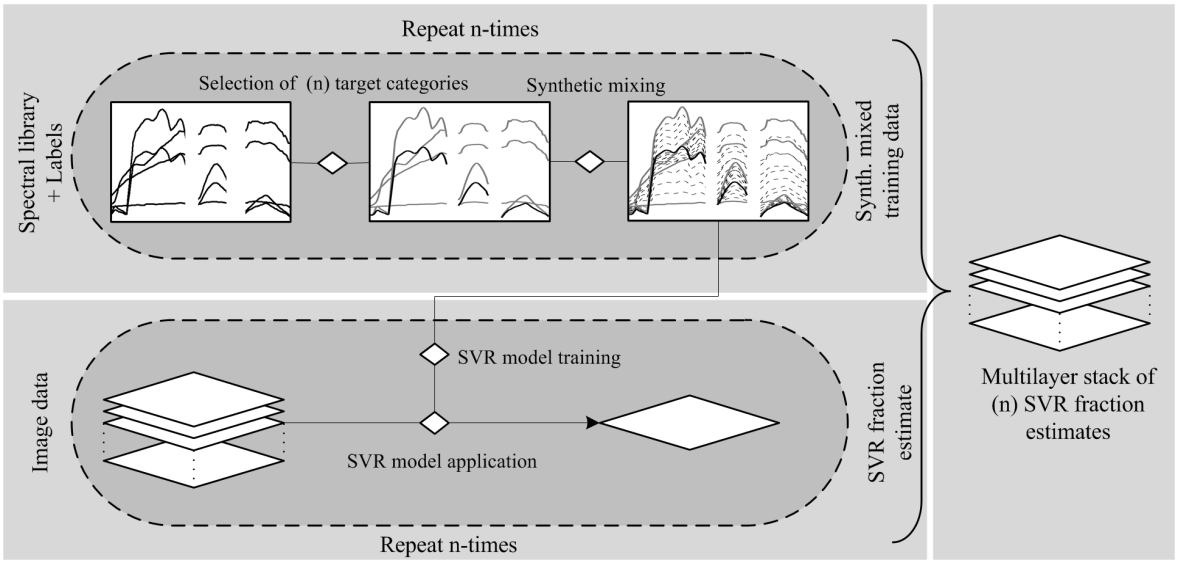


Figure A-1: Workflow for quantifying land cover using synthMix-SVR (Okujeni et al. (2013), modified).

synthMix-SVR is developed as a open source product at the Geomatics Lab of Humboldt-Universität zu Berlin. By distributing the tool, the authors hope to enlarge the number of applications and in this way learn more about its potential and weakness.

synthMix-SVR uses generic file formats (ENVI file types) as used by the EnMAP-Box 2.0. These formats are used for spectral image data as well as spectral libraries and library labels. The labeling of library spectra is not supported by synthMix-SVR and has to be carried out using the Labeling Tool of the EnMAP-Box.

The synthMix-SVR application is distributed with a test data set, consisting of a hyperspectral image subset and a corresponding spectral library. Once the data have been loaded into the EnMAP-Box, synthMix-SVR guides the user through several processing steps. This includes (i) management of input and output data, (ii) selection of target categories, (iii) definition of mixing parameters during the training data generation, and (iv) an automatized or user-defined SVR model parameterization with subsequent model application. Further optional features of synthMix-SVR enable the superposition of random noise on spectra and image post-processing to produce physically meaningful fraction maps.

2 Background

Support vector machines (SVMs) emanate from the field of machine learning and provide flexible, non-parametric and nonlinear models that are excellently suited for exploiting remote sensing image data (e.g., Foody and Mathur 2004; Melgani and Bruzzone 2004; Camps-Valls et al. 2006; Durbha et al. 2007). Detailed introductions into support vector machines and their underlying concepts can be found in Schölkopf and Smola (2002), Smola and Schölkopf (2004) or Burges (1998). While the support vector classifier (SVC) has been established as a powerful technique for the per-pixel mapping of discrete land cover classes, little attention has been paid to the use of support vector regression (SVR) for estimating sub-pixel fractions of land cover. This can be partially explained by the difficulty in finding reliable quantitative training information, i.e., pairs of spectra and associated cover fraction, needed for regression modeling. Compared to per-pixel classifiers, training signatures can hardly be labeled in the data itself or mapped in the field. A further possible solution to combine image spectra with spatially aggregated land cover information from a high resolution reference map often fails due to inaccurate co-registered data sets. In this context, the combination of SVR with synthetically mixed

training data has been demonstrated to overcome this drawback and was therefore recommended as suitable approach for sub-pixel mapping purposes (Okujeni et al. 2013).

Generation of synthetically mixed training data

The general idea behind generating synthetically mixed training data is to produce a set of multiple mixed spectra along with related mixing fractions, which can be used as training input for regression modeling of a single target land cover category (Figure A-2). A library consisting of pure material spectra that are assigned to their land cover category forms the data base. Following the description in Okujeni et al. (2013), further processing steps include:

(1) The partitioning of the library spectra into a target category and a background category (includes all remaining categories).

(2) The calculation of synthetic mixtures between each pure spectrum of the target category (with 100% mixing fraction) and each pure spectrum of the background category (with 0% mixing fraction). For simplification, linear mixing systematics is assumed. Further, the user needs to define the mixing parameters, including the mixing complexity (number of possible material spectra to be mixed) and the mixing interval (number of intermediate mixtures within the fraction range between 0 and 100%).

(3) The combination of all pure original and mixed spectra in a single spectral library. The mixing fraction of the respective target category is assigned to each spectrum.

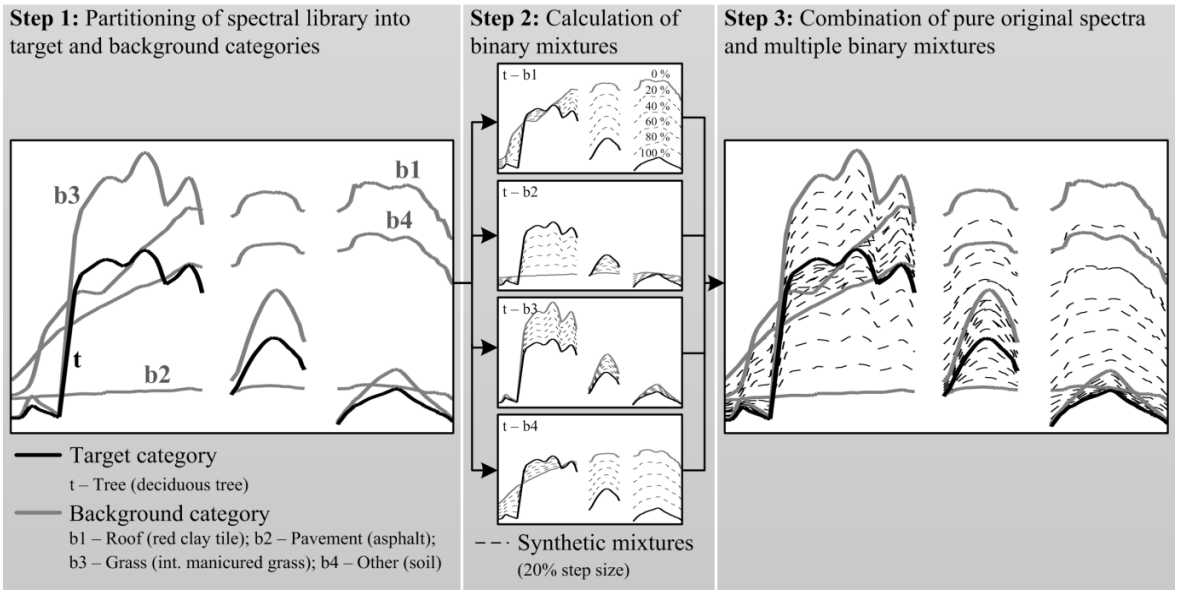


Figure A-2: Generation of synthetically mixed training data (Okujeni et al. (2013)).

synthMix-SVR was developed to generate synthetically mixed training data for multiple target categories through iterative processing. The user is requested to set the mixing interval. The current version of synthMix-SVR only supports the generation of binary mixtures between each pure spectrum of the target category and each pure spectrum of the background category. To account for environmental or instrumental errors, the user may optionally add noise to the spectral data.

Support vector regression modeling

SVR has been widely used as powerful, nonlinear technique mainly for quantifying biophysical/-chemical plant properties (Camps-Valls et al. 2006; Durbha et al. 2007; Tuia et al. 2011). In general, SVR estimates a linear dependency between pairs of n-dimensional input vectors (i.e., spectral bands) and a 1-dimensional target variable (i.e., land cover fraction of a target category) by fitting an optimal approximating hyperplane to the training data. For nonlinear problems, the training data are implicitly mapped by a kernel function into a higher dimensional space, wherein the new data distribution enables a better fitting of a linear hyperplane. The parameterization of an SVR requires the user to select the parameter(s) of a kernel function g as well as the regularization C and loss function ε parameters. Once these parameters have been selected, the optimal approximating hyperplane is found by quadratic optimization.

synthMix-SVR integrates the SVR algorithm provided by imageSVM 3.0 (available from: www.imagesvm.net). imageSVM is an IDL based tool for the SVM classification and regression analysis of remote sensing image data. imageSVM uses LIBSVM (Chang and Lin 2011) and a Gaussian kernel function during the training of the SVM. synthMix-SVR makes use of the imageSVM graphical user interface, which enables (i) the automatized or user-defined SVR model parameterization via grid search and internal validation, and (ii) the subsequent model application to derive a model prediction. Once the synthetically mixed data for the selected target categories have been generated, synthMix-SVR iteratively trains SVR models and derives fraction maps through model application to the image data.

Post-processing of fraction maps

Land cover fractions predicted by SVR cover continuous, physically meaningful fraction values between 0 and 100% through partial interpolations of the training data interval. However, improper extrapolations may also result in unrealistic fractions, i.e., negative

values (below 0%) or super-positive values (greater than 100%). Beyond, through mapping single land cover categories independently from each other, it cannot be guaranteed that the combination of all fraction maps sum to unity (100%). A comprehensive analysis and discussion is provided in Okujeni et al. (2013).

The synthMix-SVR post-processing module was designed to optionally account for unrealistic fraction values and to produce meaningful stacks of fraction maps that sum to unity.

3 User guide

3.1 Data requirements and file formats

Fraction mapping using synthMix-SVR requires a spectral image, a spectral library and a classification label file that is related to the library spectra:

(1) **Image data** are expected to be stored as ENVI Standard file. The EnMAP-Box provides functionalities for importing other file types (e.g. TIFF, ASCII).

(2) **Spectral libraries** are expected to be stored as ENVI Spectral Library file. The EnMAP-Box provides functionalities for importing other file types (e.g. ASD file, ASCII, CSV-tables).

(3) The **classification label file** is a pseudo image stored as ENVI Classification file and contains the class information of the library spectra. After importing the spectral library into the EnMAP-Box, the Labeling Tool can be used to create a corresponding classification label file.

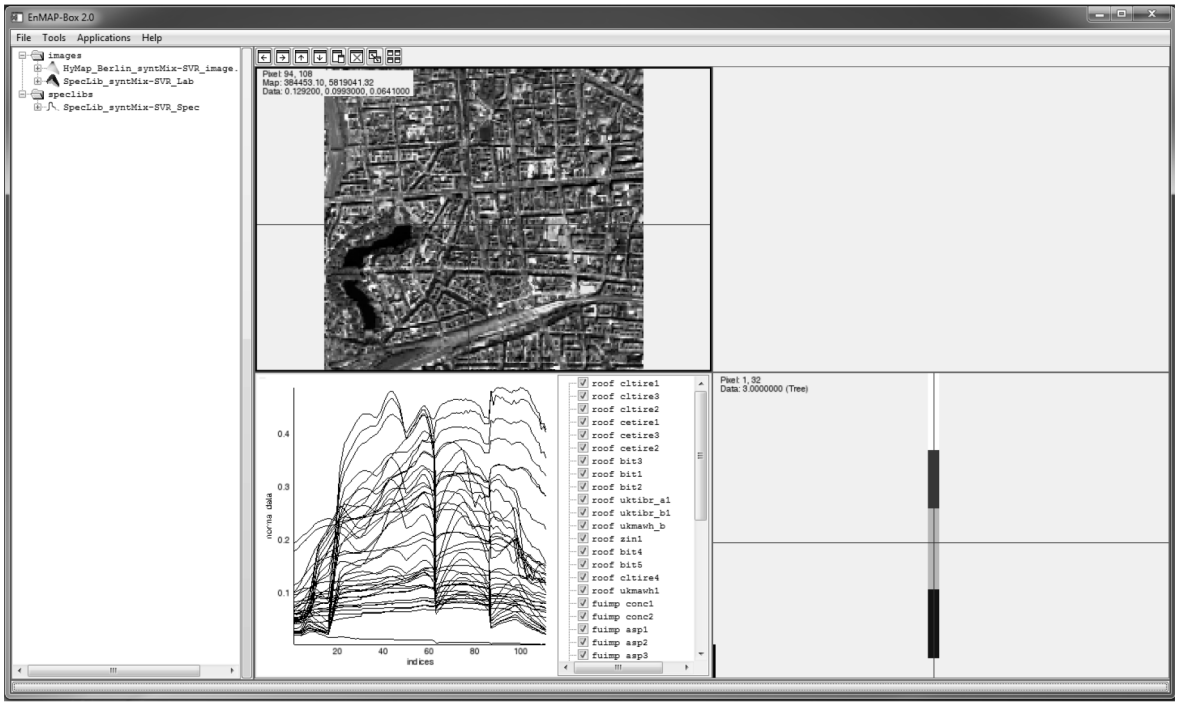
For further information on data types, the EnMAP-Box Labeling tool or the EnMAP-Box Data Format Definition please refer to (Held et al. 2014).

3.2 Getting started

After starting the **EnMAP-Box 2.0**, a test data set can be opened from the **EnMAP-Box File Menu**. It includes a subset of a hyperspectral urban scene of Berlin, Germany, (HyMap data, 9 m spatial resolution, 111 spectral bands) and a spectral library with 41 pure image spectra assigned to the four categories impervious (23 spectra), grass (5), tree (7) and other (6). More detailed information on the data set can be found in Okujeni et al. (2013) and Okujeni et al. (2014).

⇒ Select **Applications > Unmixing > synthMix-SVR > Load Test Data**. The test data set will appear in the EnMAP-Box Filelist.

⇒ Use the **drag-and-drop** functionality to display the test data set in the **EnMAP-Box View Manager**.



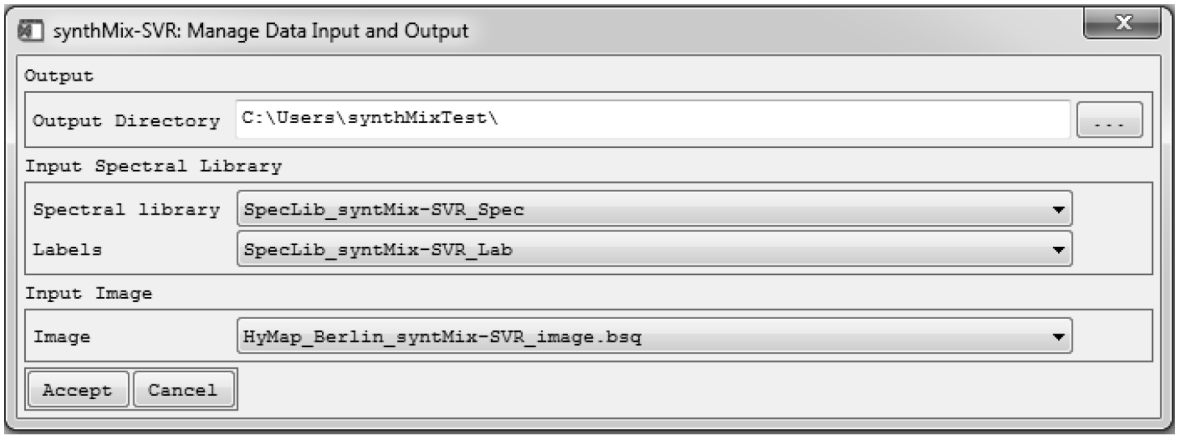
3.3 Run synthMix-SVR

3.3.1 Open synthMix-SVR

After loading the test data set, synthMix-SVR can be started from **EnMAP-Box Application Menu**.

⇒ Select **Applications > Unmixing > synthMix-SVR > synthMix-SVR**.

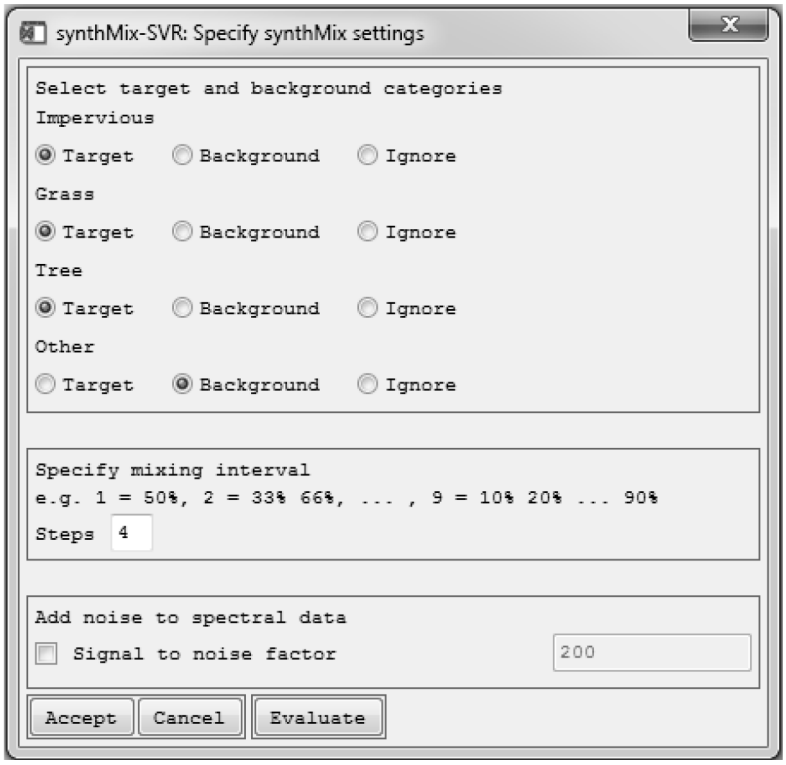
3.3.2 Manage data input and output



In the first dialog, the user is asked to specify the **input data** (spectral library, classification label and image data) and the **output data** folder where all produced data sets (training data, SVR models and fraction maps) will be stored.

⇒ Specify the **settings** (as shown above) and click **Accept** to proceed.

3.3.3 Generate synthetically mixed training data



During the generation of synthetic mixtures, the user is requested specify settings for generation synthetically mixed training data:

Select target and background categories: Displays the land cover categories according to the classification label file. The user is asked to specify the target categories (synthetic training data, SVR models and land cover fraction maps will be produced) or categories which will only be used in the background during the generation of synthetic mixtures. The ignore option may be used to exclude categories entirely from the analysis. The user must choose at least one target and one background category or two target categories.

Specify mixing interval: The mixing interval refers to the number of intermediate mixtures between 0 and 100% when mixing target category spectra (with mixing fraction of 100%) against background category spectra (with mixing fraction of 0%). The mixing interval should balance the trade-off between necessary and avoidable

number of mixtures, considering accuracy of fraction maps and computing time. In synthMix-SVR, the mixing interval is defined equally for all land cover categories and is specified via mixing steps:

Mixing step	Interval width	Output mixing fractions of target category
1	50%	0%, 50%, 100%
2	33%	0%, 33%, 66%, 100%
3	25%	0%, 20%, 40%, 60%, 89%, 100%
4	20%	0%, 25%, 50%, 75%, 100%
…	…	…
9	10%	0%, 10%, 20%, 30%, 40%, 50%, … 90%, 100%

Add noise to spectral data: Optionally, pure and synthetically mixed spectra can be imposed by noise to account for environmental or instrumental errors. The noise-degraded signal $R_{SN}(\lambda)$ is calculated according to Okin et al. (2001):

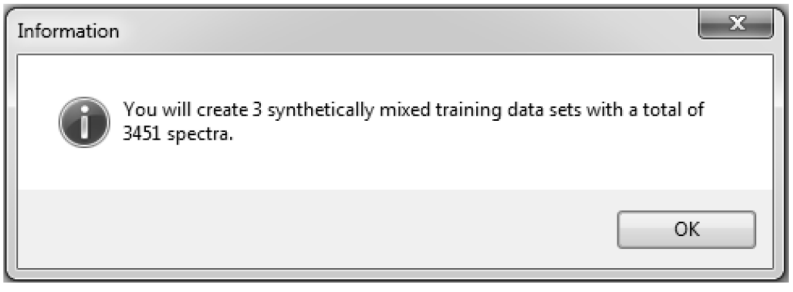
$$R_{SN}(\lambda) = R_S(\lambda)\left(1 + \frac{N(0,1)}{SNR\,(\lambda)}\right),$$

where, $R_S(\lambda)$ is the noise-free spectrum $N(0,1)$ is a random number generated from a normal distribution with a mean of 0 and standard deviation of 1 and $SNR\,(\lambda)$ is the signal-to-noise ratio.

⇒ Enter the **settings** as shown above.

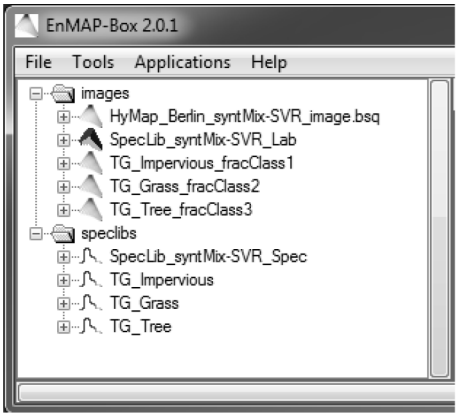
Before accepting the dialog, the user may evaluate the settings in order to get an overview of how many synthetically mixed training data sets and spectra are created. Note that the SVR processing time increases with increasing input data.

⇒ Click **Evaluate.**



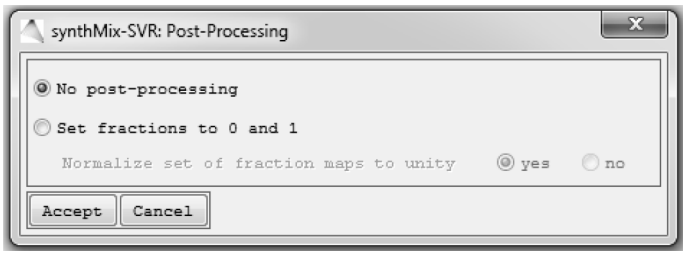
⇒ Click **OK** to get back to previous dialog. Adjust settings if necessary and click **Accept** to proceed.

The generated **synthetically mixed training data sets**, i.e., pure original and synthetically mixed spectra as well as related mixing fractions, for the selected target categories will be created, saved in the defined output folder and opened in the **EnMAP-Box File Menu**.



3.3.4 Specify post-processing

The following dialog allows the user to optionally specify post-processing options that will be carried out once the fraction maps are produced.



No post-processing: Fraction maps will not be post-processed.

Set fractions to 0 and 1: Negative (below 0%) and super-positive (above 100%) fraction values will be set to 0 and 100%.

Normalize set of fraction maps: Weights the estimated fractions of a target category relative to the sum of estimated fractions of all target categories on a per-pixel basis. This way all target category fraction maps sum to unity (i.e., 100%).

⇒ Specify settings and click **Accept** to proceed.

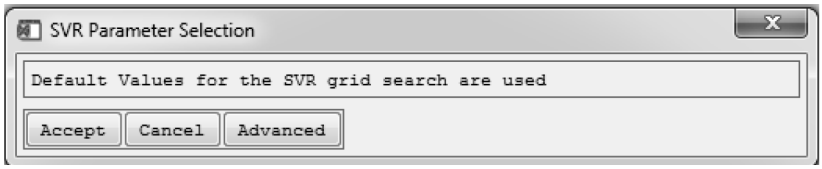
Note that all files, including original and the post-processed fraction maps will be saved in the output folder when post-processing options are selected.

3.3.5 SVR modeling

SVR modeling using synthMix-SVR is based on imageSVM, which allows the user to parameterize SVR models based on default values or advanced settings prior to applying

the model to the image data. SVR parameters *g* and *C*, depend on the data range and distribution and are thus case specific. A common strategy to search for adequate values for *g* and *C* is a two-dimensional grid search with internal validation. This strategy is implemented in imageSVM. For tuning the loss function parameter *ε* we use a efficient search heuristic proposed by Rabe et al. (2013)

SVR modeling using default values



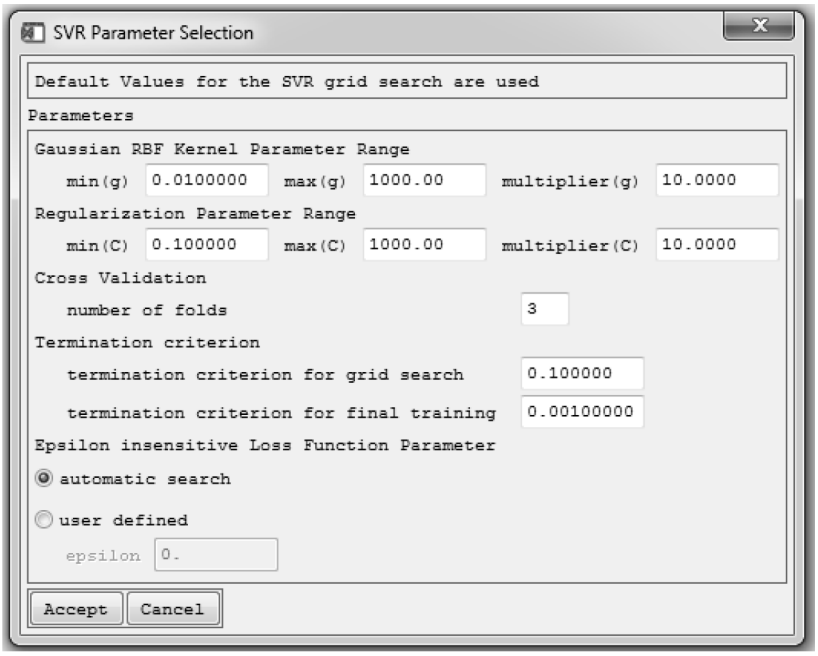
Default values for the grid search will be used to find ideal parameter values for the Gaussian kernel *g*, regularization *C* and loss function *ε* during model parameterization for all target categories. In most cases, default values already lead to high accuracies.

⇒ Click **Accept** to proceed with SVR modeling using default values. The default values can be examined by clicking **Advanced**.

SVR modeling using advanced settings

The user may change the default values for the grid search or specify user defined parameters for the Gaussian kernel *g*, regularization *C* and loss function *ε* during model parameterization for all target categories.

⇒ Click **Advanced** to continue with the advanced settings. The SVC parameterization dialog is now expanded.



The user is now allowed to modify the grid search and may select:

min (*g*/*C*), max (*g*/*C*): Minimum and maximum values that define the range of the grid (*g* and *C* dimension).

Multiplier (*g*/*C*): Specifies the step size of the grid.

Cross Validation: The accuracy of results during the grid search is monitored by n-fold cross validation on the training data.

Termination criterion: Tolerance of termination criterion during grid search or final model training.

Epsilon loss insensitive Loss Function Parameter: Apply efficient search heuristic or enter user defined values.

⇒ Adjust the settings if wanted and click **Accept** to proceed with SVR modeling using advanced settings.

synthMix-SVR will start with the **SVR model training** and **model application** for each target category by iterative processing. Corresponding SVR models and fraction maps will be saved in the output folder and opened in the **EnMAP-Box Filelist**.

The user may manually check the **SVR parameters** (e.g., selected parameter sets during grid search, training error, etc.,) by opening the SVR models in the **imageSVM Regression Application Menu**.

⇒ **Applications > Regression > imageSVM Regression > View SVR Parameters**

For further information on SVR modeling using imageSVM, please refer to the imageSVM Regression Manual (van der Linden et al. 2014).

3.4 synthMix-SVR results

All generated files are saved in the specified output folder and opened in the EnMAP-Box:

(1) Original input spectral library, library labels and spectral image

(2) Synthetically mixed training data of the selected target categories

(3) SVR models of the selected target categories

(4) Single fraction map of each target category

(5) Stack of all fraction maps in a multilayer image

(6) (Optional) Stack of fraction maps with values between 0 and 100% in a multilayer image.

(7) (Optional) Stack of normalized fraction maps in a multilayer image.

An output HTML-report will be opened in the standard browser for providing general information about the synthMix-SVR process.

The user may display fraction maps in a RGB color composite and establish a link to the original spectral image

Use the **drag-and-drop** functionality to display the test data set in the **EnMAP-Box View Manager**. The illustration below shows a composite of grass (R), tree (G) and impervious (B).

3.5 References

Burges, C.J.C. (1998). A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.

Camps-Valls, G., Bruzzone, L., Rojo-Alvarez, J.L., & Melgani, F. (2006). Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geoscience and Remote Sensing Letters*, 3, 339-343.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2.

Durbha, S.S., King, R.L., & Younan, N.H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, 107, 348-361.

Foody, G.M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. IEEE *Transactions on Geoscience and Remote Sensing*, 42, 1335-1343.

Held, M., Hostert, P., Jakimow, B., van der Linden, S., Okujeni, A., Rabe, A., Suess, S., & Wirth, F. (2014). *EnMAP-Box Documentation Center*, Version 2.0.2. Germany: Humboldt-Universität zu Berlin.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1778-1790.

Okin, G.S., Roberts, D.A., Murray, B., & Okin, W.J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, 77, 212-225.

Okujeni, A., van der Linden, S., Jakimow, B., Rabe, A., Verrelst, J., & Hostert, P. (2014). A Comparison of Advanced Regression Algorithms for Quantifying Urban Land Cover. *Remote Sensing*, 6, 6324-6346.

Okujeni, A., van der Linden, S., Tits, L., Somers, B., & Hostert, P. (2013). Support vector regression and synthetically mixed training data for quantifying urban land cover. *Remote Sensing of Environment*, 137, 184-197.

Rabe, A., Jakimow, B., van der Linden, S., Okujeni, A., Suess, S., Leitao, P.J., & Hostert, P. (2013). Simplifying Support Vector Regression Parameterisation by Heuristic Search for Optimal e-Loss. *8th EARSeL Workshop of Special Interest Group in Imaging Spectroscopy*. Nantes, France.

Schölkopf, B., & Smola, A.J. (2002). *Learning with Kernels - Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, Massachusetts: MIT Press.

Smola, A.J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14, 199-222.

Tuia, D., Verrelst, J., Alonso, L., Perez-Cruz, F., & Camps-Valls, G. (2011). Multioutput support vector regression for remote sensing biophysical parameter estimation. *IEEE Geoscience and Remote Sensing Letters*, 8, 804-808.

van der Linden, S., Rabe, A., Held, A., Wirth, F., Suess, S., Okujeni, A., & Hostert, P. (2014). *imageSVM Regression, Manual for Application: imageSVM version 3.0*. Humboldt-Universität zu Berlin, Germany.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbstständig und ohne Verwendung unerlaubter Hilfe angefertigt zu haben. Die aus fremden Quellen direkt oder indirekt übernommenen Inhalte sind als solche kenntlich gemacht. Die Dissertation wird erstmalig und nur an der Humboldt-Universität zu Berlin eingereicht. Weiterhin erkläre ich, nicht bereits einen Doktortitel im Fach Geographie zu besitzen. Die dem Verfahren zu Grunde liegende Promotionsordnung ist mir bekannt.

Akpona Okujeni

Berlin, den 20.08.2014